# Unbiased Adaptive Estimations of the Fourth-Order Cumulant for Real Random Zero-Mean Signal

Iaroslav V. Blagouchine and Eric Moreau, *Senior Member, IEEE*

***Abstract*—In this paper, a consistent efficient estimator of the fourth-order cumulant for real discrete-time random i.i.d. (at least up to order 8) zero-mean signal is proposed, in both, batch and adaptive versions. In batch version, the proposed estimator is not only consistent, but also unbiased and efficient. The systematical theoretical and experimental studies with comparisons between the proposed estimator and three other estimators of the fourth-order cumulant (the natural or the traditional one, the trivial unbiased estimator for the known power case and the fourth $k$-statistics), are undertaken, for both, normal and uniform processes. Then, the adaptive versions of the estimators (all, except the fourth $k$-statistics), are given and studied in detail. The convergence in mean and the convergence in mean square analyses are performed for them, first theoretically, then empirically. Finally, the whole set of analyses carried out for both batch and adaptive versions shows that from many points of view the proposed estimator is interesting for use in versatile signal processing applications, especially in real-time and short-term ones.**

***Index Terms*—Adaptive estimation, bias, consistency, convergence in mean, convergence in mean square, cumulant, estimation, estimator, higher moments, higher order statistics (HOS), $k$-statistics, mean square error (MSE), random signals, recursive method, semi-invariant, stochastic processes, variance.**

## I. INTRODUCTION

THE use of the higher-order statistics in signal processing is nowadays an ordinary procedure. The third- and fourth-order moments and cumulants are especially of great interest, since they found many practical applications: blind source separation problems [1]–[3], including for the MIMO systems [4], which by the way found many applications in wireless communications (e.g., Wi-Fi routers with multiple antennæ), identification of FIR systems [5], [6], speech stream and voicing detections [7]–[9], speech recognition [10]–[12], general speech processing [13], and many others. Usually, *cumulants*, also known as *cumulative moment functions*, *semi-invariants*, or *half-invariants* [14]–[18], are more often used in applications, and the moments have generally only an auxiliary function. Unlike the moments, cumulants cannot be calculated directly from the density probability function $p(x)$, but only via the characteristic function or the moments. Since the moments and the cumulants can be easily expressed in terms of each other [15], [17]–[19], the latter are often computed via the moments.

Manuscript received October 06, 2008; accepted March 26, 2009. First published April 21, 2009; current version published August 12, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Antonio Napolitano.

The authors are with the Telecommunication Department, ISITV of the University of Toulon, F-83162, Valette du Var (Toulon), France (e-mail: iaroslav.blagouchine@univ-tln.fr; moreau@univ-tln.fr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2009.2021453

Let us consider a real random discrete signal $x_i \equiv x(i)$, $i \in \mathbb{Z}$, where $i$ is discrete time, i.e., number of current sample. Let us consider in addition this signal is zero-mean and i.i.d. up to order 8; in other words, the samples of $x_i$ are identically distributed ($x_i$ is stationary) and independent up to order 8. Furthermore, it is also supposed that all its raw moments up to order 8 exist. The latter are denoted for simplicity by $\mathrm{E}[x^p] \equiv \mathrm{E}[x_i^p]$, $p \in \mathbb{N}^*$, where $\mathrm{E}[\cdot]$ is the operator of the mathematical expectation.

The moment-based definition of the fourth-order cumulant $\kappa_4$ for such a signal $x_i$ is given by

$$\kappa_4 = \mathrm{E}[x^4] - 3\mathrm{E}^2[x^2] \tag{1}$$

where the raw moments $\mathrm{E}[x^2]$ and $\mathrm{E}[x^4]$ are calculated from the density probability function $p(x)$ defining the distribution of the samples $x_i$. In practice, in many cases, on the one hand, we do not know exactly the density probability function $p(x)$, and on the another hand, we do not have an access to all samples of $x_i$. Thus, in practice, the raw moments $\mathrm{E}[x^p]$, $p \in \mathbb{N}^*$, cannot be calculated directly, and consequently, neither the cumulant in question. In these cases, we can make an estimation of this cumulant from a single realization of $n$ samples of $x_i$. The classical or natural estimator of this cumulant is given by [20]

$$\hat{\kappa}_{4,\mathrm{nat}} = \frac{1}{n}\sum_{i=1}^{n} x_i^4 - \frac{3}{n^2}\left(\sum_{i=1}^{n} x_i^2\right)^2. \tag{2}$$

This estimator is called "natural," because the unknown moments are simply, or naturally, replaced by the sample ones. This estimator was subject to numerous studies in literature over the past 60 years, and many authors use this estimator for their works. Notwithstanding its simplicity and the fact that this estimator is very often used in practice, there is no reason that it is the best estimator of the fourth-order cumulant. An estimator is characterized by three fundamental properties: consistency (absence of bias in probability or asymptotically, i.e., when $n \to \infty$), bias and efficiency, expressed in terms of mean square error of the estimator [21]–[23]. The estimator (2) is consistent, but it is biased and has nonzero variance, which is, in turn, closely related to its mean square error (MSE).

In fact, the general problem of the unbiased estimation of cumulants is an important statistical problem having potentially many versatile applications in the field of signal processing. The solution of this problem is known over the past 80 years under the name of the so-called $k$-statistics [14], [18], [24], [25]. These statistics, denoted by $k_p$, $p \in \mathbb{N}^*$, the mean values of which are unconditionally equal to the $p$th cumulants $\mathrm{E}[k_p] = \kappa_p$, were worked out in a quite general way by Fisher at the end of the twenties of the XXth century [14], [18]. They provide the unbiased estimations of the cumulants for the very general case, in which all the previous moments (i.e., moments of order less or equal than that of cumulant) are considered unknown. These unknown moments are replaced by the weighted symmetric product sums, and the final expressions may be expressed in terms of sample means. For instance, the fourth $k$-statistics is given by

$$k_4 = \frac{n^2\left\{(n+1)m_4 - 3(n-1)m_2^2\right\}}{(n-1)(n-2)(n-3)} \tag{3}$$

where $m_2$ and $m_4$ are the second and the fourth sample central moments

$$m_p = \frac{1}{n}\sum_{i=1}^{n}(x_i - m_1')^p$$
$$m_1' = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad p\in\mathbb{N}^*$$

[14], [18], [25]. However, in many signal processing problems some moments are explicitly known, and consequently, the estimation could be more accurate. The most frequent case is that of the strictly centered processes; e.g., speech signals, audio signals, almost all telecommunication signals. In these cases, one can suppose that the fourth $k$-statistics, that additionally estimates the mean value, becomes less efficient, and, therefore, it may be not the best choice for such situations. Thus, we decided to use this *a priori* information in order to build a precise and efficient estimator of the fourth-order cumulant. We will then compare it to the other aforementionned estimators of the fourth-order cumulant, as well as to the trivial unbiased estimator for the known parent variance $\mathrm{E}[x^2]$ case (i.e., known power case, which can be often considered, e.g., for frequency and phase modulation signals), in both, batch and adaptive versions.

## II. Batch Versions of the Estimators

### A. Construction of the Unbiased Estimators

*1) Unbiased Estimator for the Unknown Variance Case:* The main problem of the natural estimator (2) is the second term in the above difference. The first term of (2) is the unbiased estimator of $\mathrm{E}[x^4]$, but the second term (quadratic term) introduces a bias to the estimation of the term $\mathrm{E}^2[x^2]$. In fact, if we estimated $\mathrm{E}^2[x^2]$ via the following estimator:

$$\frac{1}{n^2}\left(\sum_{i=1}^{n} x_i^2\right)^2 \tag{4}$$

the bias $b$ would be

$$\begin{aligned} b &= \mathrm{E}\left[\frac{1}{n^2}\left(\sum_{i=1}^{n} x_i^2\right)^2\right] - \mathrm{E}^2[x^2] \\ &= \mathrm{E}\left[\frac{1}{n^2}\sum_{i} x_i^4 + \frac{1}{n^2}\sum_{\substack{i,j\\ i\neq j}} x_i^2 x_j^2\right] - \mathrm{E}^2[x^2] \\ &= \frac{1}{n}\Big(\mathrm{E}[x^4] - \mathrm{E}^2[x^2]\Big). \end{aligned} \tag{5}$$

Note that in the above calculation, we used the following property:

$$\mathrm{E}[x_i^p x_j^q] = \begin{cases} \mathrm{E}[x_i^p]\,\mathrm{E}[x_j^q] & i\neq j \\ \mathrm{E}[x_i^{p+q}] & i=j \end{cases} \qquad p,q\in\mathbb{N}^* \tag{6}$$

provided that $p+q\leqslant 8$, and which holds because $x_i$ is assumed to be an i.i.d. up to order 8 signal. Strictly speaking, here, we only need the fulfilment of the property (6) for $p$ and $q$ such that $p+q\leqslant 4$, but in further calculations concerning variance, the more strict condition $p+q\leqslant 8$ is required. Therefore, estimator (4) is biased, but still consistent because

$$\lim_{n\to\infty} b = 0. \tag{7}$$

Consequently, the natural estimator (2) is also biased, and its bias is

$$b = \frac{3}{n}\Big(\mathrm{E}^2[x^2] - \mathrm{E}[x^4]\Big) \tag{8}$$

but it still remains consistent because (7) holds as well.

Before proceed with the construction of the unbiased estimators of the fourth-order cumulant, we will just recall one important theorem from theory of estimations: *if one can found an unbiased symmetric polynomial estimator of the* $p$th *cumulant, this estimator is the unique unbiased estimator of this cumulant* [18]. Hence, the unbiased estimators of the cumulants we will construct are unique.

From previous analysis, one can notice that by choosing properly the coefficients before two terms in the right part of (2), one can compensate the bias introduced by the quadratic term, and therefore, make the total bias vanishes. Let us introduce now the following estimator of the cumulant $\kappa_4$:

$$\hat{\kappa}_{4,\mathrm{unk}} = \alpha\sum_{i=1}^{n} x_i^4 - \beta\left(\sum_{i=1}^{n} x_i^2\right)^2. \tag{9}$$

Its bias is

$$b = \mathrm{E}[\hat{\kappa}_{4,\mathrm{unk}}] - \kappa_4 = (\alpha n - \beta n - 1)\mathrm{E}[x^4] - (\beta n(n-1) - 3)\mathrm{E}^2[x^2]$$

Hence, if we want the bias to be zero, we must choose the coefficients $\alpha$ and $\beta$ to be equal to

$$\alpha = \frac{n+2}{n(n-1)}, \quad \beta = \frac{3}{n(n-1)}. \tag{10}$$

The resulting unbiased estimator of the fourth cumulant is

$$\hat{\kappa}_{4,\mathrm{unk}} = \frac{n+2}{n(n-1)}\sum_{i=1}^{n} x_i^4 - \frac{3}{n(n-1)}\left(\sum_{i=1}^{n} x_i^2\right)^2. \tag{11}$$

As a matter of fact, note that it is also possible to arrive to this expression [as well as to obtain the further formulas (17), (18), and (19) in Section II-C1], by using the aforementioned Fisher's method employed for the obtaining of $k$-statistics.

*2) Unbiased Estimator for the Known Variance Case:* By considering the parent variance $\mathrm{E}[x^2]$ (or power of the signal $x_i$), denoted by $\sigma^2$, a known fixed value, the cumulant to estimate becomes

$$\kappa_4 = \mathrm{E}[x^4] - 3\sigma^4$$

Obviously, the unbiased estimation of this kind of cumulant is

$$\hat{\kappa}_{4,\mathrm{kno}} = \frac{1}{n}\sum_{i=1}^{n} x_i^4 - 3\sigma^4. \tag{12}$$

In our work, this trivial estimator will be also used, mainly for comparison purposes.

### B. Efficiencies of the Estimators and Their Comparisons

Besides consistency and bias, another important property of the estimator is its efficiency, which is usually expressed in terms of its MSE [21]–[23]. For the statistics in question, the latter is defined as

$$\mathrm{MSE} \equiv \mathrm{E}\Big[\big(\hat{\kappa}_4 - \kappa_4\big)^2\Big] = \mathrm{E}[\hat{\kappa}_4^2] - 2\kappa_4\mathrm{E}[\hat{\kappa}_4] + \kappa_4^2. \tag{13}$$

The MSE is often written in terms of variance and bias

$$\mathrm{var}\,[\hat{\kappa}_4] + b^2 = \mathrm{MSE} \tag{14}$$

where the variance is

$$\mathrm{var}\,[\hat{\kappa}_4] \equiv \mathrm{E}\Big[\big(\hat{\kappa}_4 - \mathrm{E}[\hat{\kappa}_4]\big)^2\Big] = \mathrm{E}[\hat{\kappa}_4^2] - \mathrm{E}^2[\hat{\kappa}_4]. \tag{15}$$

Thus, since the MSE takes account of both, bias and variance, it is usually employed as the index of efficiency of an estimator. Note by the way that we do not call it "the index of performance," because the latter may be defined differently, depending on the concrete application and aim. For instance, one can consider a different cost functions composed of bias, variance, MSE, entropy, likelihood, posterior expected value of a loss function (Bayesian estimation), etc., weighed by the corresponding coefficients. By properly choosing the nature of each term and each weight, one can emphasize the desired characteristics of an estimator and this function can be called "the index of performance."

From latter equations, we understand that the main problem of the calculation of MSE is actually reduced to the calculation of the term $\mathrm{E}[\hat{\kappa}_4^2]$, the term $\kappa_4^2$ being given by (1), and $\mathrm{E}[\hat{\kappa}_4]$ was calculated before, during the calculation of bias. The direct calculation of the term $\mathrm{E}[\hat{\kappa}_4^2]$, is often long, that is why we present it in detail only once for $\hat{\kappa}_{4,\mathrm{unk}}$, in Appendix A; the calculation of this term for other estimators being almost analogous to the presented one.

We will now compare the efficiencies of the four aforementioned estimators: $\hat{\kappa}_{4,\mathrm{unk}}$ given by (11), $\hat{\kappa}_{4,\mathrm{kno}}$ given by (12), $\hat{\kappa}_{4,\mathrm{nat}}$ given by (2) and $k_4$ given by (3). Since the most frequent and important distributions in signal processing are normal and uniform (e.g., distribution of the speech signal samples are nearly Gaussian, almost any natural noise, including thermal noise, is also normal, quantization error in analog-to-digital converters is uniform, messages emitted from a discrete source having the maximum entropy are distributed uniformly, etc.), we will study only the processes described by these two distributions. The even raw moments of the Gaussian and uniform distributions with zero mean and given variance $\sigma^2$, denoted $N(0,\sigma)$ and $U(0,\sigma)$, are, respectively

$$\mathrm{E}[x^{2u}] = \frac{2^u\Gamma\big(u+\frac{1}{2}\big)}{\sqrt{\pi}}\sigma^{2u} = \sigma^{2u}\prod_{l=1}^{u}(2l-1), \qquad u = 1,2,\ldots$$

and

$$\mathrm{E}[x^{2u}] = \frac{3^u\sigma^{2u}}{2u+1}, \qquad u = 1,2,\ldots$$

where $\Gamma(\cdot)$ is the gamma function, also known as Euler's integral of the second kind. As to the odd moments $\mathrm{E}[x^{2u+1}]$, they are all null for both distributions because of the symmetry.

The general formula for the MSE of the estimator for the unknown variance case (11) is given by (see Appendix A for details):

$$\begin{aligned}\mathrm{E}\Big[\big(\hat{\kappa}_{4,\mathrm{unk}} - \kappa_4\big)^2\Big] &= \frac{1}{n}\mathrm{E}[x^8] - \frac{12}{n}\mathrm{E}[x^6]\mathrm{E}[x^2] - \frac{n-19}{n(n-1)}\mathrm{E}^2[x^4] \\ &\quad + \frac{48n-84}{n(n-1)}\mathrm{E}[x^4]\mathrm{E}^2[x^2] - \frac{36n-54}{n(n-1)}\mathrm{E}^4[x^2]. \end{aligned} \tag{16}$$

So, it tends asymptotically to zero when $n \to \infty$. For the aforementioned normal and uniform processes, the MSE becomes

$$\begin{cases} \mathrm{MSE}_{\mathrm{unk}}\,[N(0,\sigma)] = \dfrac{24(n+2)}{n(n-1)}\sigma^8 \\ \mathrm{MSE}_{\mathrm{unk}}\,[U(0,\sigma)] = \dfrac{288(6n+1)}{175n(n-1)}\sigma^8. \end{cases}$$

For the trivial estimator for the known variance case (12), the MSE is

$$\mathrm{E}\Big[\big(\hat{\kappa}_{4,\mathrm{kno}} - \kappa_4\big)^2\Big] = \frac{1}{n}\Big(\mathrm{E}[x^8] - \mathrm{E}^2[x^4]\Big).$$

In this case, the MSE also tends asymptotically to zero. For the particular cases we face, it reads

$$\begin{cases} \mathrm{MSE}_{\mathrm{kno}}\,[N(0,\sigma)] = 96\dfrac{\sigma^8}{n} \\ \mathrm{MSE}_{\mathrm{kno}}\,[U(0,\sigma)] = 5.76\dfrac{\sigma^8}{n}. \end{cases}$$

As to the natural estimator (2), since it is biased, the calculation of the MSE is slightly more complicated. First, we calculate its variance according to (15), and then use (14) in order to calculate the corresponding MSE. For the variance, we first need the mean value of the natural estimator

$$\mathrm{E}[\hat{\kappa}_{4,\mathrm{nat}}] = \frac{(n-3)}{n}\mathrm{E}[x^4] - \frac{3(n-1)}{n}\mathrm{E}^2[x^2].$$

Then, after the calculation of the term $\mathrm{E}[\hat{\kappa}_{4,\mathrm{nat}}^2]$, which is quite similar to that of $\mathrm{E}[\hat{\kappa}_{4,\mathrm{unk}}^2]$ performed in the Appendix A, we obtain the variance of the natural estimator

$$\begin{aligned}\mathrm{var}\,[\hat{\kappa}_{4,\mathrm{nat}}] &= \frac{(n-3)^2}{n^3}\mathrm{E}[x^8] - \frac{12(n-1)(n-3)}{n^3}\mathrm{E}[x^6]\mathrm{E}[x^2] \\ &\quad - \frac{n^2-24n+27}{n^3}\mathrm{E}^2[x^4] + \frac{12(n-1)(4n-9)}{n^3}\mathrm{E}[x^4]\mathrm{E}^2[x^2] \\ &\quad - \frac{18(n-1)(2n-3)}{n^3}\mathrm{E}^4[x^2]\end{aligned}$$

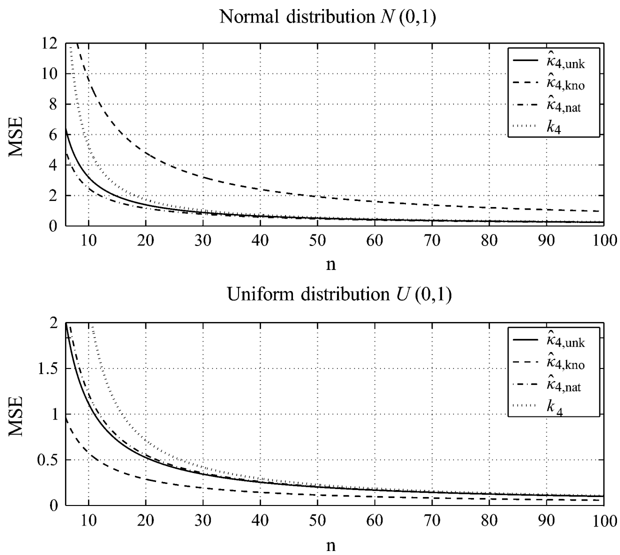


Fig. 1. The evolution of the resulting MSE for the normal and uniform distributions.

and finally, its MSE

$$\mathrm{E}\Big[\big(\hat{\kappa}_{4,\mathrm{nat}}-\kappa_4\big)^2\Big] = \frac{(n-3)^2}{n^3}\mathrm{E}[x^8] - \frac{12(n-1)(n-3)}{n^3}\mathrm{E}[x^6]\mathrm{E}[x^2] - \frac{n^2-33n+27}{n^3}\mathrm{E}^2[x^4] + \frac{6(8n^2-29n+18)}{n^3}\mathrm{E}[x^4]\mathrm{E}^2[x^2] - \frac{9(n-2)(4n-3)}{n^3}\mathrm{E}^4[x^2].$$

Hence, once again, both variance and MSE tend asymptotically to zero. For the considered uniform and normal cases, the MSE becomes

$$\begin{cases} \mathrm{MSE}_{\mathrm{nat}}\,[N(0,\sigma)] = \dfrac{12(2n^2-3n+36)}{n^3}\sigma^8 \\ \mathrm{MSE}_{\mathrm{nat}}\,[U(0,\sigma)] = \dfrac{144(12n^2+29n-6)}{175n^3}\sigma^8. \end{cases}$$

Finally, as to the MSE of the fourth $k$-statistics, the corresponding expression can be found in [14], [18] or [25] (in first two references it is designated by special notation $\kappa(4^2) \equiv \kappa_2[k_4] = \mathrm{var}[k_4]$), and it tends asymptotically to zero as well. For the particular normal and uniform cases with which we deal, it, respectively, gives

$$\begin{cases} \mathrm{MSE}_{\mathrm{kst}}\,[N(0,\sigma)] = \dfrac{24n(n+1)}{(n-1)(n-2)(n-3)}\sigma^8 \\ \mathrm{MSE}_{\mathrm{kst}}\,[U(0,\sigma)] = \dfrac{288(6n^3+5n^2+n+315)}{175n(n-1)(n-2)(n-3)}\sigma^8. \end{cases}$$

The behavior of the resulting MSE for all four estimators is given in the Fig. 1. As can be ascertained from these graphics, the efficiency of the unbiased estimator $\hat{\kappa}_{4,\mathrm{unk}}$ is almost the same as that of the natural estimator $\hat{\kappa}_{4,\mathrm{nat}}$ for both, normal and uniform distributions,[1] but since the estimator $\hat{\kappa}_{4,\mathrm{unk}}$ is unbiased, it is more precise. As to the estimator for the known variance case $\hat{\kappa}_{4,\mathrm{kno}}$, its behavior is strongly influenced by the distribution of the initial data, and depending on the conditions, it can be efficient or inefficient. Finally, as to the fourth $k$-statistics $k_4$, it is less efficient than the proposed estimator $\hat{\kappa}_{4,\mathrm{unk}}$ and than the natural one $\hat{\kappa}_{4,\mathrm{nat}}$, in both uniform and normal scenarios.

Taking into account the latter fact, and that the analysis of the fourth $k$-statistics in adaptive version would be too tedious,[2] the fourth $k$-statistics will not be considered for the further studies devoted to the adaptive versions of the estimators.

### C. Extensions to the Higher-Order and to the Cross-Cumulant Cases

*1) Higher-Order Extensions:* our method, based on the replacement of the unknown raw moments by the sample ones with modified coefficients, may be used for the construction of the unbiased estimators for other higher-order cumulants. In particular, the unbiased estimator of the third-order cumulant $\kappa_3[x] = \mathrm{E}[x^3]$ is simply the third sample mean (i.e., in this case, the unique coefficient $n^{-1}$ remains unchanged)

$$\hat{\kappa}_3 = \frac{1}{n}s_3, \qquad \text{where} \quad s_r \equiv \sum_{i=1}^{n} x_i^r, \qquad r \in \mathbb{N}^* \tag{17}$$

that of the fifth-order cumulant $\kappa_5[x] = \mathrm{E}[x^5] - 10\mathrm{E}[x^3]\mathrm{E}[x^2]$ is

$$\hat{\kappa}_5 = \frac{n+9}{n(n-1)}s_5 - \frac{10}{n(n-1)}s_3 s_2 \tag{18}$$

that of the sixth-order cumulant

$$\kappa_6[x] = \mathrm{E}[x^6] - 15\mathrm{E}[x^4]\mathrm{E}[x^2] - 10\mathrm{E}^2[x^3] + 30\mathrm{E}^3[x^2]$$

is

$$\hat{\kappa}_6 = \frac{1}{n(n-1)(n-2)}\Big\{(n^2+22n+12)s_6 - 15(n+4)s_4 s_2 - 10(n-2)s_3^2 + 30 s_2^3\Big\} \tag{19}$$

and so on. In fact, the possibility to generalize the method is simply due to the fact, that the cumulants are expressed in terms of different raw moment products, which includes all possible homogeneous combinations of these moments (this follows directly from the well-known moment-to-cumulant formula [17]–[19]); these raw moments are then replaced by the sample ones with variable coefficients, and the bias, due to the crossed terms into the products of sample moments, can be precisely compensated by the modifications of these coefficients.

*2) Cross-Cumulant Extensions:* We finally would like to briefly discuss the possible extensions of the method to the

[1]By the way, the MSE for $n \gg 1$ are in first approximation equal for the estimators $\hat{\kappa}_{4,\mathrm{unk}}, \hat{\kappa}_{4,\mathrm{nat}}$ and $k_4$: $\mathrm{MSE}\,[N(0,\sigma)] = (24n^{-1} + O(n^{-2}))\,\sigma^8$, and $\mathrm{MSE}\,[U(0,\sigma)] \approx (9.87n^{-1} + O(n^{-2}))\,\sigma^8$, while the MSE for $\hat{\kappa}_{4,\mathrm{kno}}$ in first approximation differ from previous ones for both, normal and uniform distributions.

[2]An approximate volume of such calculations can be estimated basing on the Appendices B and C, where we, respectively, calculated the mean and the MSE of the adaptive estimator for the unknown variance case, and this statistics does not imply even a half of the auxiliary statistics that are needed for the fourth $k$-statistics.

cross-cumulants (called also joint cumulants), which are constantly receiving growing interest in versatile signal processing applications. Let us assume we have four real random discrete-time zero-mean i.i.d. signals $u_i$, $x_i$, $y_i$ and $z_i$. If these processes are dependent (otherwise, such cross-cumulants are null), we may again apply our method to obtain the unbiased estimators. If, for example, we wished to estimate without bias the third-order cross-cumulant $\kappa_3[x,y,z] = \mathrm{E}[x,y,z]$, [17]–[19], such an estimator would be the following one:

$$\hat{\kappa}_3[x,y,z] = \frac{1}{n}\sum_{i=1}^{n} x_i y_i z_i.$$

As to the unbiased estimation of the fourth-order cross-cumulant, defined as [17]–[19]

$$\kappa_4[u,x,y,z] = \mathrm{E}[u,x,y,z] - \mathrm{E}[u,x]\mathrm{E}[y,z] - \mathrm{E}[u,y]\mathrm{E}[x,z] - \mathrm{E}[u,z]\mathrm{E}[x,y]$$

its unbiased estimator may be given by

$$\hat{\kappa}_4[u,x,y,z] = \alpha\sum_{i=1}^{n} u_i x_i y_i z_i - \beta\left(\sum_{i=1}^{n} u_i x_i\right)\left(\sum_{l=1}^{n} y_l z_l\right) - \gamma\left(\sum_{i=1}^{n} u_i y_i\right)\left(\sum_{l=1}^{n} x_l z_l\right) - \delta\left(\sum_{i=1}^{n} u_i z_i\right)\left(\sum_{l=1}^{n} x_l y_l\right)$$

where the coefficients $\alpha, \beta, \gamma, \delta$ are found, as previously, in such way that the bias vanishes (obviously, they depend on the concrete dependencies between $u_i, x_i, y_i, z_i$). In analogous manner, the extensions to higher-order cross-cumulants may be also obtained basing on the above-mentioned moment-to-cumulant formula, but it should be taken into account that the more the order of the cumulant and the more the complexity of the dependency between the processes, the more long and complicated the determination of the coefficients.

## III. Adaptive Versions of the Estimators

The adaptive estimation is based on a recurrent estimation over time index $n$. The aim of the adaptive estimation is to be able to calculate the new value of estimation, by using the previous value of the estimation and the new observation. Basing on the recurrent relationship for the estimator, an adaptive step size $\mu$, which is usually real and chosen in the range $0 < \mu \ll 1$, is then introduced as: $\mu = 1/n$.

### A. Construction of the Estimators and Convergence in Mean

*1) Unknown Variance Case:* If we denote the estimator (11) by $\hat{\kappa}_{4,n,\mathrm{unk}}$, where $n$ is the size of frame we dispose for each single estimation, one can see that its adaptive version can be written as

$$\hat{\kappa}_{4,n+1,\mathrm{unk}} = \frac{(1-\mu)(1+3\mu)}{(1+\mu)(1+2\mu)}\hat{\kappa}_{4,n,\mathrm{unk}} + \frac{3\mu}{(1+\mu)(1+2\mu)} \cdot \hat{\sigma}_n^4 - \frac{6\mu}{1+\mu}x_{n+1}^2\hat{\sigma}_n^2 + \frac{\mu}{1+\mu}x_{n+1}^4 \tag{20}$$

where by $\hat{\sigma}_n$ we denoted the square root of the unbiased estimation of variance for the zero-mean case made from a frame of $n$ samples

$$\hat{\sigma}_n = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}. \tag{21}$$

Thus, unfortunately, the adaptive estimation of the fourth-order cumulant also involves at each stage the additional adaptive estimation of $\hat{\sigma}_n$

$$\hat{\sigma}_n^2 = (1-\mu)\hat{\sigma}_{n-1}^2 + \mu x_n^2. \tag{22}$$

In practice, in order to initialize the algorithm, besides the choice of the adaptive step size $\mu$, we must also choose the initial values $\hat{\sigma}_1^2$ and $\hat{\kappa}_{4,2,\mathrm{unk}}$. For instance, for the forthcoming empirical tests, we chose to initialize and $\hat{\sigma}_1^2$ at 0, and $\hat{\kappa}_{4,2,\mathrm{unk}}$ at $-0.6$ (mean between the values of the fourth semi-invariant for the normal and uniform distributions of unit standard deviation), but this choice is not mandatory and these values can be any constants in the reasonable interval.

We now study the convergence in mean of the adaptive algorithm. The analysis of the convergence in mean of this estimator is a quite long procedure, that is why we put it into Appendix B, and here we give only the final result [see (23) shown at the bottom of the page], with $f_0$ defined in (42). Then, by calculating the limit of the latter, that is fortunately easy to do since its 3 terms of 4 vanish when $n$ tends to infinity, we obtain

$$\lim_{n\to\infty}\mathrm{E}[\hat{\kappa}_{4,n+1,\mathrm{unk}}] = \frac{2}{(5\mu+1)(2-\mu)} \Big\{(-\mu^2+3\mu+1)\mathrm{E}[x^4] - 3(-2\mu^2+4\mu+1)\mathrm{E}^2[x^2]\Big\} \tag{24}$$

with the domain of convergence defined as intersection of $(5-\sqrt{33})/2 < \mu < -1/5$ and $0 < \mu < (5+\sqrt{33})/2$, due to

$$\begin{aligned}\mathrm{E}[\hat{\kappa}_{4,n+1,\mathrm{unk}}] = {} & f_0^{n-1}\hat{\kappa}_{4,2,\mathrm{unk}} + \frac{(1-\mu)^{2n-2} - f_0^{n-1}}{(1-\mu)^2 - f_0}\cdot\frac{3\mu(1-\mu)^2}{(1+\mu)(1+2\mu)} \\ & \cdot\left\{\hat{\sigma}_1^4 - \frac{\mu}{2-\mu}\mathrm{E}[x^4] - \frac{2(1-\mu)}{2-\mu}\mathrm{E}^2[x^2] - 2\mathrm{E}[x^2]\Big(\hat{\sigma}_1^2 - \mathrm{E}[x^2]\Big)\right\} \\ & + \frac{\mu(1-f_0^{n-1})}{(1+\mu)(1-f_0)}\left\{\mathrm{E}[x^4] - 6\mathrm{E}^2[x^2] + \frac{3}{1+2\mu}\left(\frac{\mu}{2-\mu}\mathrm{E}[x^4] + \frac{2(1-\mu)}{2-\mu}\mathrm{E}^2[x^2]\right)\right\} \\ & - \frac{(1-\mu)^{n-1} - f_0^{n-1}}{1-\mu-f_0}\cdot\frac{12\mu^2(1-\mu)}{(1+\mu)(1+2\mu)}\mathrm{E}[x^2]\Big(\hat{\sigma}_1^2 - \mathrm{E}[x^2]\Big)\end{aligned} \tag{23}$$

the first term in the estimator (20), with $0 < \mu < 2$, due to the first term in the auxiliary estimator (22); so, finally, the common domain of convergence is $0 < \mu < 2$.

Hence, this estimator has a bias, but it vanishes for small step sizes $\mu$, i.e., when $\mu \to +0$. Also, one can easily notice that there is another interesting case when the bias also vanishes, and it is also in the range of convergence: $\mu = 1$. We have

$$\lim_{\substack{n\to\infty \\ \mu\to+0}} \mathrm{E}[\hat{\kappa}_{4,n+1,\mathrm{unk}}] = \lim_{\substack{n\to\infty \\ \mu\to 1}} \mathrm{E}[\hat{\kappa}_{4,n+1,\mathrm{unk}}] = \mathrm{E}[x^4] - 3\mathrm{E}^2[x^2].$$

This second case may also be important because for small step size $\mu$ the convergence is slow, but for $\mu = 1$ it is very fast. Moreover, from (24) it follows also that there are no more unbiased cases for our estimator, since the second degree equation has only two roots. Thus, using the terminology of the adaptive estimation literature, the estimator $\hat{\kappa}_{4,n+1,\mathrm{unk}}$ is said to be quasi-convergent in mean; i.e., there is no convergence in mean in a classical sense, since the mean does not tend to $\kappa_4$ when $n \to \infty$ for any $\mu$ in the domain of convergence, but the quasi-convergence in mean, because the mean tends to $\kappa_4$ when $n \to \infty$ and $\mu \to +0$.

*2) Known Variance Case:* This is again a quite simple case for which the adaptive version can be written as

$$\hat{\kappa}_{4,n+1,\mathrm{kno}} = \frac{1}{1+\mu}\hat{\kappa}_{4,n,\mathrm{kno}} + \frac{\mu}{1+\mu}x_{n+1}^4 - \frac{3\mu}{1+\mu}\sigma^4 \quad (25)$$

that follows directly from (12).

Now, we perform the analysis of convergence of this adaptive algorithm. By proceeding in the analogous way to that of the adaptive estimator of variance, we obtain first

$$\hat{\kappa}_{4,n+1,\mathrm{kno}} = \frac{1}{(1+\mu)^n}\hat{\kappa}_{4,1,\mathrm{kno}} + \frac{\mu}{1+\mu}\sum_{k=0}^{n-1}\frac{x_{n-k+1}^4}{(1+\mu)^k} - 3\sigma^4\left[1 - \frac{1}{(1+\mu)^n}\right].$$

Then, by calculating its mean value:

$$\mathrm{E}[\hat{\kappa}_{4,n+1,\mathrm{kno}}] = \frac{1}{(1+\mu)^n}\hat{\kappa}_{4,1,\mathrm{kno}} + \left[1 - \frac{1}{(1+\mu)^n}\right]\mathrm{E}[x^4] - 3\sigma^4\left[1 - \frac{1}{(1+\mu)^n}\right]. \quad (26)$$

And finally, by evaluating the limit of the last mean

$$\lim_{n\to\infty}\mathrm{E}[\hat{\kappa}_{4,n+1,\mathrm{kno}}] = \begin{cases} \mathrm{E}[x^4] - 3\sigma^4 & \text{if } \mu < -2 \cup \mu > 0 \\ \text{diverges} & \text{otherwise.} \end{cases} \quad (27)$$

So this adaptive estimator is convergent, but unlike the previous ones, it is always unbiased. In other words, this estimator is convergent in mean in a classical sense. Note, by the way, that this estimator is the unique one which converges for negative and for great step sizes $\mu$.

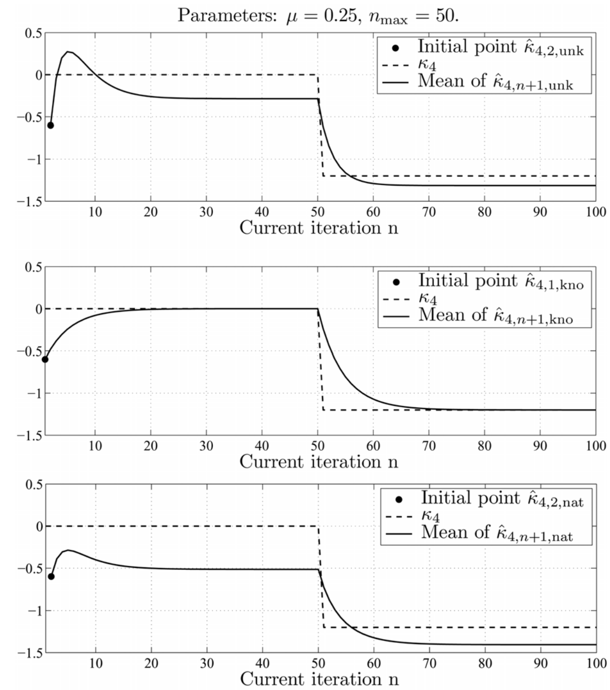


Fig. 2. Theoretical behavior of the means of the estimators. By $n_{\max}$ we denoted the maximum number of iterations $n$ for both $N(0,1)$ and $U(0,1)$ scenarios.

*3) Natural Estimator:* For the natural estimator (2), the adaptive version can be written as follows:

$$\hat{\kappa}_{4,n+1,\mathrm{nat}} = \frac{1}{1+\mu}\hat{\kappa}_{4,n,\mathrm{nat}} + \frac{3\mu}{(1+\mu)^2}\hat{\sigma}_n^4 - \frac{6\mu}{(1+\mu)^2}x_{n+1}^2\hat{\sigma}_n^2 + \frac{\mu(1-2\mu)}{(1+\mu)^2}x_{n+1}^4. \quad (28)$$

Again, the adaptive estimation involve the calculation of the statistics $\hat{\sigma}_n$, which must be calculated from its previous value via (22).

Analogously to the adaptive convergence in mean of the algorithm for the unknown variance, we first have

$$\begin{aligned}\mathrm{E}[\hat{\kappa}_{4,n+1,\mathrm{nat}}] &= \frac{\hat{\kappa}_{4,2,\mathrm{nat}}}{(1+\mu)^{n-1}} + \left\{1 - \frac{1}{\left[(1+\mu)(1-\mu)^2\right]^{n-1}}\right\} \\ &\cdot \frac{3\mu(1-\mu)^{2n}}{(1+\mu)\left[(1+\mu)(1-\mu)^2 - 1\right]} \\ &\cdot \left\{\hat{\sigma}_1^4 - \frac{\mu\mathrm{E}[x^4]}{2-\mu} - \frac{2(1-\mu)}{2-\mu}\mathrm{E}^2[x^2] \right. \\ &\left. - 2\mathrm{E}[x^2]\left(\hat{\sigma}_1^2 - \mathrm{E}[x^2]\right)\right\} + \frac{(1+\mu)^{n-1} - 1}{(1+\mu)^n} \\ &\cdot \left\{(1-2\mu)\mathrm{E}[x^4] - 6\mathrm{E}^2[x^2] \right. \\ &\left. + 3\left(\frac{\mu}{2-\mu}\mathrm{E}[x^4] + \frac{2(1-\mu)}{2-\mu}\mathrm{E}^2\left[x^2\right]\right)\right\} \quad (29)\end{aligned}$$

and then

$$\lim_{n\to\infty} \mathrm{E}[\hat{\kappa}_{4,n+1,\mathrm{nat}}] = \frac{2}{(1+\mu)(2-\mu)} \cdot\Big\{(\mu^2-\mu+1)\mathrm{E}[x^4] - 3\mathrm{E}^2[x^2]\Big\} \quad (30)$$

provided that $0 < \mu < 2$. Strangely, this estimator becomes unbiased exactly for the same values of $\mu$ as it was for the unknown variance case estimator: when it tends to $+0$, and when $\mu = 1$. Furthermore, these unbiased cases are unique for this estimator as well, since (30) is a second degree equation for $\mu$. Thus, the natural estimator is not convergent in mean in a classical sense, but quasi-convergent in mean.

*4) Comparisons:* In the first place, we present the graphs of the theoretical behavior of the means of three considered estimators of the fourth-order cumulant for given $\mu$, implemented according to (23), (26), and (29). First half of data is considered being distributed according to the law $N(0,1)$, and the second half, to the law $U(0,1)$; the adaptive step size $\mu$ is set to 0.25. Note that we deliberately choose $\mu$ not very small in order to avoid a relatively banal behavior[3] of the statistics. First of all, we note that the behavior of our estimator for the unknown variance case and that of the natural estimator is quite similar, both having a maximum that occurs in the beginning when $n = 5$, and after reaching it, they begin to converge to their final limit value, defined by the corresponding bias. Moreover, it is quite well visible that the latter is almost twice greater for the natural estimator. Namely, the bias of the estimator for the unknown variance case is equal to $-0.286$ for the normal case, and $-0.114$, for the uniform case, while the bias of the natural estimator is about $-0.514$ and $-0.206$, respectively [see (24) and (30)]. On the contrary, the estimator for the known variance case has a qualitatively different behavior. It does not have any maximum in the beginning and converges directly to the true value of cumulant, since it has no bias. As to the rate of convergence, it is practically the same for all the statistics, but in other cases, it can differ, namely, in the case $\mu = 1$ that we come to study in the experimental section of our work, Section IV.

In the second place, we illustrate the behavior of the asymptotic bias as a function of $\mu$, according to (24), (27), and (30), see Fig. 3.

The dependence of the asymptotic bias on the step size $\mu$ shows the asymptotic bias is always smaller for the estimator $\hat{\kappa}_{4,n+1,\mathrm{unk}}$ than that for the natural one $\hat{\kappa}_{4,n+1,\mathrm{nat}}$ in both, normal and uniform scenarios. For both estimators and in both scenarios, it is slightly negative for $0 < \mu < 1$, it is positive tending to infinity for $1 < \mu < 2$ and it is null for $\mu = +0$ and $\mu = 1$.

### B. Convergence in Mean Square

*1) Unknown Variance Case and Natural Estimator:* The convergence in mean square of these algorithms, necessary for the corresponding MSE analysis, is usually performed in order to find out the efficiencies of estimators. The latter analysis is even more long and tedious than that of its bias, because for both estimators $\hat{\kappa}_{4,n+1,\mathrm{unk}}$ and $\hat{\kappa}_{4,n+1,\mathrm{nat}}$ it implies a lot of additional calculations related to four auxiliary estimators $\hat{\sigma}_n^2$, $\hat{\sigma}_n^4$, $\hat{\sigma}_n^6$, $\hat{\sigma}_n^8$, which are in addition all self-implicative (e.g., $\hat{\sigma}_n^8$ requires $\hat{\sigma}_n^6$, $\hat{\sigma}_n^4$ and $\hat{\sigma}_n^2$), as well as, to four mixed estimators $\hat{\kappa}_{4,n,\mathrm{unk}}\hat{\sigma}_n^2$, $\hat{\kappa}_{4,n,\mathrm{unk}}\hat{\sigma}_n^4$, $\hat{\kappa}_{4,n,\mathrm{nat}}\hat{\sigma}_n^2$ and $\hat{\kappa}_{4,n,\mathrm{nat}}\hat{\sigma}_n^4$, which are also self-implicative. Thus, this analysis was performed asymptotically and put into Appendix C. From (44), it follows that both estimators are quasi-convergent in mean square (i.e., the MSE $\to 0$ when $n \to \infty$ and $\mu \to +0$).

[3]We mean, we do not want to take $\mu \approx +0$, which is usual in adaptive estimations, because it would lead to the unbiasedness and to the entirely similar behavior of both estimators $\hat{\kappa}_{4,n+1,\mathrm{unk}}$ and $\hat{\kappa}_{4,n+1,\mathrm{nat}}$.

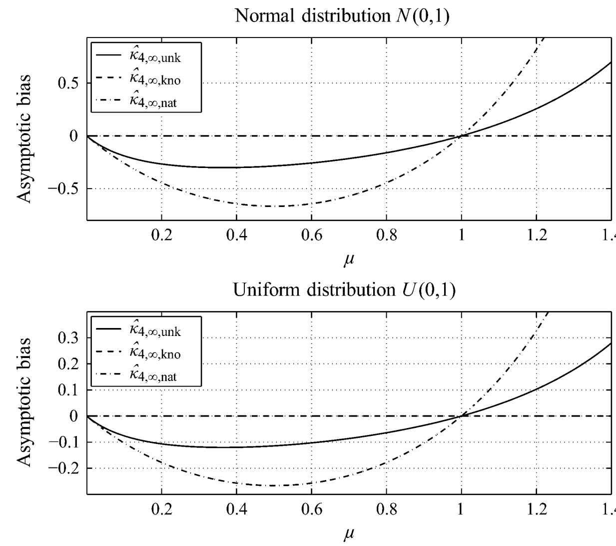


Fig. 3. Dependence of the asymptotic bias on the step size $\mu$.

*2) Known Variance Case:* The asymptotic MSE of this adaptive estimator is much easier to calculate than the previous ones. Basing on the method described in the Appendix C, and under assumption of the convergence in mean (27), we find

$$\lim_{n\to\infty} \mathrm{E}\Big[\big(\hat{\kappa}_{4,n+1,\mathrm{kno}} - \kappa_4\big)^2\Big] = \frac{\mu}{2+\mu}\Big(\mathrm{E}[x^8] - \mathrm{E}^2[x^4]\Big). \quad (31)$$

It tends to zero for $\mu \to +0$ and for $\mu \to -\infty$, and, thus, the estimator is also quasi-convergent in mean square.

*3) Comparisons:* We compare the behaviors of the asymptotic MSE of the three considered estimators as a function of $\mu$, according to (44) and (31), see Fig. 4. This dependence shows for small and average step sizes $\mu$, the asymptotic MSE of the estimator for the unknown variance case and of the natural one are almost the same for the normal process, and it is slightly worse for average step sizes $\mu$ for the uniform process. As to the trivial estimator for the known variance case, its behavior is again strongly influenced by the initial distribution of the random signal $x_i$.

Finally, it should be noted that for all three estimators this second unbiased case, which gives, as we shall see later, a very fast convergence rate and that could be very attractive perspective, gives also greater MSE.

## IV. Experimental Study of the Estimators

In this section, we will experimentally study the considered estimators. To carry out the experiments the following procedure is performed. We generate $d$ different random realizations

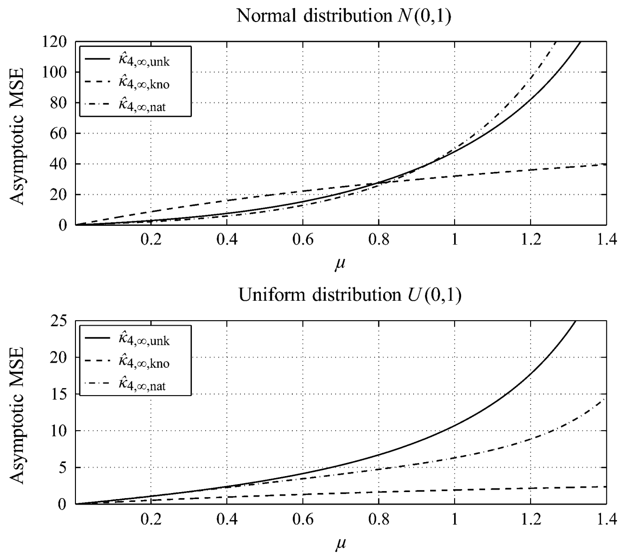


Fig. 4. Dependence of the asymptotic MSE on the step size $\mu$.

of signal $x_i$, each realization being of length $n$ (i.e., $d$ frames of $n$ samples each, giving $d\times n$ samples in all, are generated). The realizations are independent. Mathematically, the whole set of $d$ realizations (or frames) of length $n$ can be written in vectorial form as

$$\{\boldsymbol{x}_i\}_{i=1}^{n}, \qquad \forall \boldsymbol{x}_i \in \mathbb{R}^d \tag{32}$$

where $d$ is the dimension of the column vector $\boldsymbol{x}_i$. This can be also view as a $d\times n$ matrix $\mathbf{X}$, whose elements were taken from signal $x_i$. Then, for each realization (frame of length $n$), a single value of the corresponding statistics is calculated; e.g., for the natural estimator (2), one realization of $x_i$ of $n$ samples (i.e., one frame) gives one value of the estimator $\hat{\kappa}_4$. Since we have $d$ different realizations ($d$ frames) of signal $x_i$, $d$ different values of $\hat{\kappa}_4$ will be calculated. Since we are mainly interested in mean behavior of the estimator, from these $d$ different values of $\hat{\kappa}_4$, denoted by a vector $\hat{\boldsymbol{\kappa}}_4$, we will calculate the mean $\mathrm{E}[\hat{\boldsymbol{\kappa}}_4]$; the latter being obviously estimated via the sample or arithmetic mean.

Note that sometimes, especially in adaptive estimations, one prefers to show a single realization of statistics, rather that its mean (i.e., in practice, $\mathrm{E}[\hat{\boldsymbol{\kappa}}_4]$ with $d=1$, instead of $\mathrm{E}[\hat{\boldsymbol{\kappa}}_4]$ with great $d$). In particular, this can be interesting when the variance of an estimator is small enough (e.g., when $n\to\infty$ for the batch estimators, or when $\mu\to+0$ for the adaptive ones), and a single realization of the estimator has great chances to be close to the mean value of estimator, and hence, in some sense, it "replaces" the mean value. Thus, in general cases, when $n$ is not very great, or $\mu$ is not very small, we prefer to show the mean and the MSE, while for the special cases, we report also a single realization of the statistics.

### A. Batch Versions of the Estimators

The experimental studies of the batch versions of estimators are divided in two parts: static power case and dynamic power case. In both cases, we generated 500 random experiences (each experience being $d$ realizations of $n$ samples of signal $x_i$): the first 250, according to the normal law with zero mean, and the second 250, according to the uniform law with zero mean; parameters $d$ and $n$ being fixed to 800 and 10, respectively.[4] In static power case, the standard deviation of the initial data $x_i$ is set to 1 for both normal and uniform distributions. In dynamic power case, the standard deviation of the initial data $x_i$ depends on the number of current experience, denoted by $k$, according to the amplitude modulation law

$$\sigma(k) = \sigma_0\big(1 + a\sin\omega_0 k\big), \quad \omega_0 = \frac{2\pi}{k_{\max}} \tag{33}$$

with $\sigma_0 = 1$, depth of modulation $a = 0.5$, $k_{\max} = 250$ and $k = 1, 2, \ldots, k_{\max}$, for both normal and uniform cases.

The comparison of four considered in theoretical part statistics $\hat{\kappa}_{4,\mathrm{unk}}$, $\hat{\kappa}_{4,\mathrm{kno}}$, $\hat{\kappa}_{4,\mathrm{nat}}$ and $k_4$ is shown in the Fig. 5 (static power case) and Fig. 6 (dynamic power case).

For the static power case of unit power, the theoretical (or true) value of the fourth-order cumulant is equal 0 for the normal law, and to $-1.2$ for the uniform one. From Fig. 5, we can observe the mean values of the estimators $\hat{\kappa}_{4,\mathrm{unk}}$, $\hat{\kappa}_{4,\mathrm{kno}}$ and $k_4$ converge to the theoretical value of the fourth-order cumulant for both, normal and uniform distributions, while the estimator $\hat{\kappa}_{4,\mathrm{nat}}$ is strongly biased. Theoretically, according to (8), its bias is $b = -6/n = -0.6$ for the case $N(0,1)$, and is $b = -2.4/n = -0.24$ for the case $U(0,1)$, and that is precisely what we observe from Fig. 5. As to the MSE, in the Fig. 5 we find the same features and values as those obtained theoretically at the end of Section II-B and shown in Fig. 1 for parameter $n = 10$.

The dynamic power case, intended to better represent the reality (e.g., the real received signals may be strongly influenced by the fluctuating physical properties of the channels of propagation, e.g., Rayleigh fading, speech signals also have nonconstant power, etc.), is shown in the Fig. 6. Theoretical value of the fourth-order cumulant in the dynamic power case depends on the evolution of the standard deviation of the initial data: it remains still zero for normally distributed data, but it becomes equal to $-1.2\sigma^4(k)$ for the uniformly distributed data, and, therefore, depends also on the current experience $k$. From the Fig. 6, it follows the estimators $\hat{\kappa}_{4,\mathrm{unk}}$, $\hat{\kappa}_{4,\mathrm{kno}}$, and $k_4$ are again without bias, in both, static and dynamic power cases. In turn, the natural estimator has always a nonzero bias, and the latter becomes especially important in the dynamic power case for the normally distributed data. One can also notice that this bias becomes greater when the initial standard deviation reaches its maximal value 1.5. This is not an accident and actually follows directly from (8); more precisely, we have the bias

$$b\big[N(0,\sigma)\big] = -\frac{6}{n}\sigma^4\bigg|_{\sigma=1.5} \approx -3.04. \tag{34}$$

As to the MSE, it is again in accordance with the conclusions of Section II-B and Fig. 1.

[4]We deliberately chose not great $n$ in order to better observe the differences between different estimators (bias and MSE).

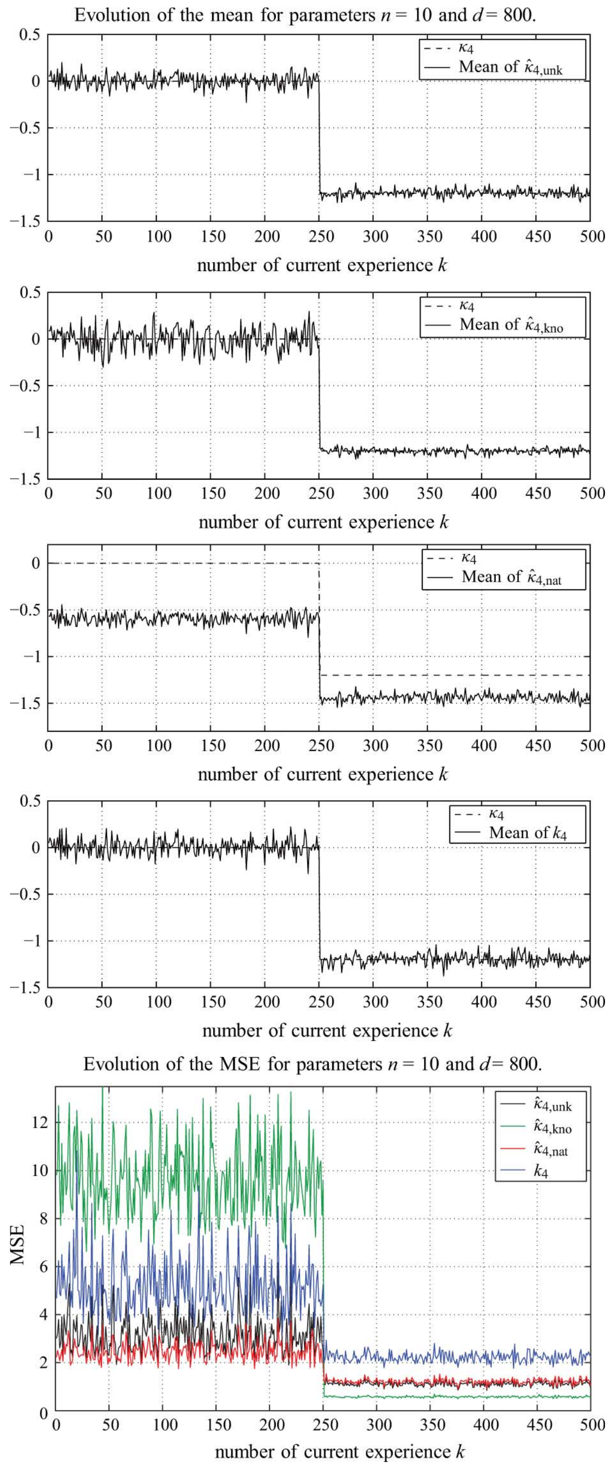


Fig. 5. Static power case: experimental means and MSE of the four considered estimators with the true theoretical value of cumulant $\kappa_4$.

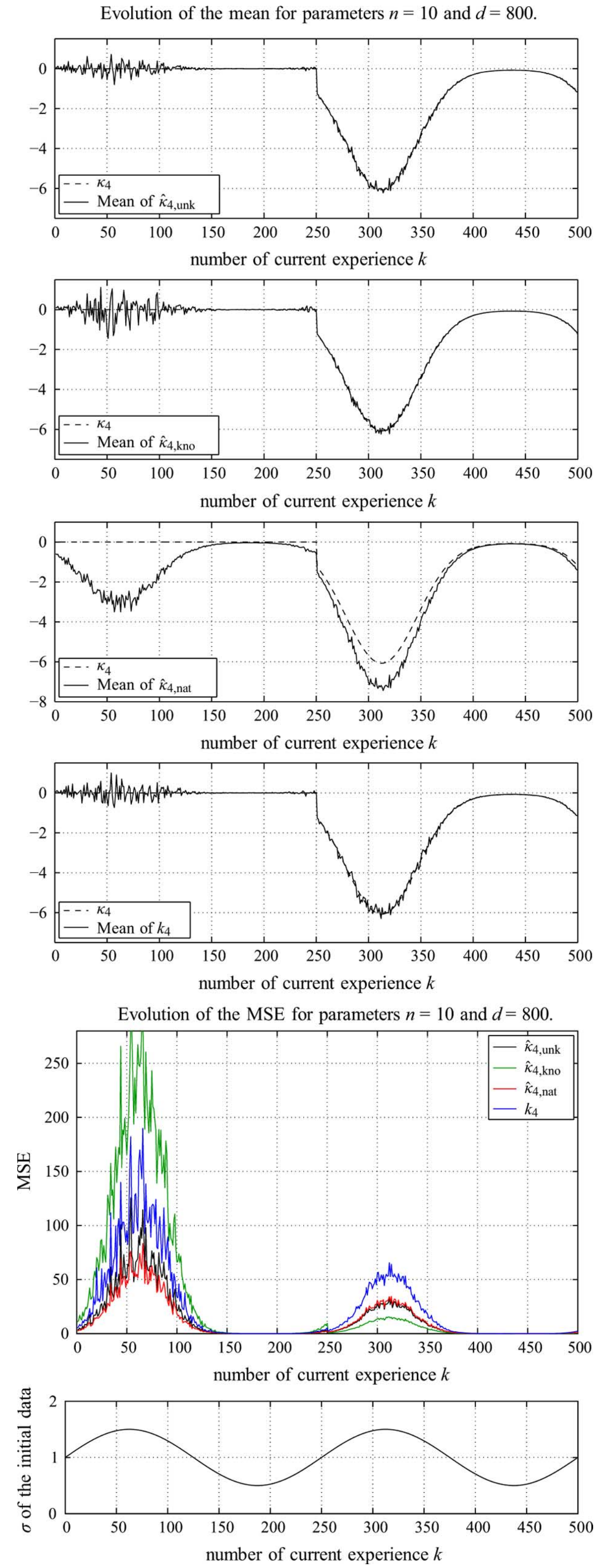


Fig. 6. Dynamic power case: experimental means and MSE of the four considered estimators, with the true theoretical value of cumulant $\kappa_4$, as well as, the evolution of the standard deviation $\sigma(k)$ of the initial random signal $x_i$.

In conclusion note, that we considered here not a banal case, when $n = 10$. If we took a more banal case (i.e., almost an asymptotic case), when $n$ is great, e.g., $n = 10000$, the behavior of the estimators would be quite different. On the one hand, the variance tending to zero, would permit to visualize the single realization of the estimators $\hat{\kappa}_4$, instead of its mean $\mathrm{E}[\hat{\kappa}_4]$. On the other hand, great $n$ would give a practically null bias for natural estimator, and we could not observe the differences between the estimators $\hat{\kappa}_{4,\mathrm{unk}}$, $\hat{\kappa}_{4,\mathrm{nat}}$ and $k_4$, because, in addition, their MSE are in first approximation equal. More generally, one can

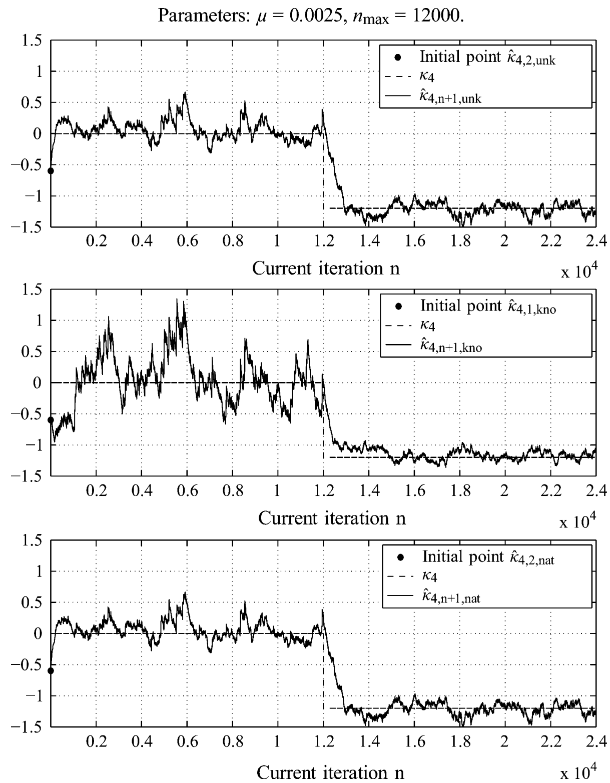


Fig. 7. Static power case: single realizations of estimators.

easily notice that for $n \gg 1$, estimators $\hat{\kappa}_{4,\mathrm{unk}}$ (11) and $\hat{\kappa}_{4,\mathrm{nat}}$ (2) become in first approximation equal

$$\frac{(n+2)}{n(n-1)} \sim \frac{1}{n} \quad \text{and} \quad \frac{3}{n(n-1)} \sim \frac{3}{n^2}.$$

Similar reasoning may be applied to $k_4$ (3), but one has also to take into account the first and third sample raw moments on which $k_4$ depends (explicitly on the first moment, and implicitly on the third one[5]); the latter both tend to zero since the studied processes are zero-mean and symmetric. Thus, the proposed estimator $\hat{\kappa}_{4,\mathrm{unk}}$ may be especially advantageous and useful in short-term and real-time signal processing applications.

### B. Adaptive Versions of the Estimators

*1) First Unbiased Case* $0 < \mu < 1$*:* As previously, in all experiments, the first half of data was distributed according to the law $N(0,1)$, and the second half, according to the law $U(0,1)$; $n_{\max}$ is the number of maximum iterations of adaptive estimators for each law.

First of all, we give a results for a single realization of estimators, Fig. 7. All estimations vary a lot locally, but their global behavior is quite stable. The step size $\mu$ was chosen equal to 0.0025, that guarantees the bias of the statistics practically null and the small MSE; Fig. 8 illustrates the mean and MSE behaviors. On the other hand, the convergence is attempted after 2000 iterations approximately. One can also notice that the estimators $\hat{\kappa}_{4,n+1,\mathrm{unk}}$ and $\hat{\kappa}_{4,n+1,\mathrm{nat}}$ behave almost equally, from both, bias and MSE points of view. This result is normal, because for small step sizes $\mu$, the initial formulas for both estimators (20) and (28) become in first approximation the same; namely, one may easily verify by using Maclaurin series, that for small $\mu$

$$\frac{(1-\mu)(1+3\mu)}{(1+\mu)(1+2\mu)} = 1 - \mu + O(\mu^2)$$

and

$$\frac{1}{1+\mu} = 1 - \mu + O(\mu^2).$$

Analogous comparisons may be done for other coefficients in (20) and (28). Also, the asymptotic equality of both adaptive

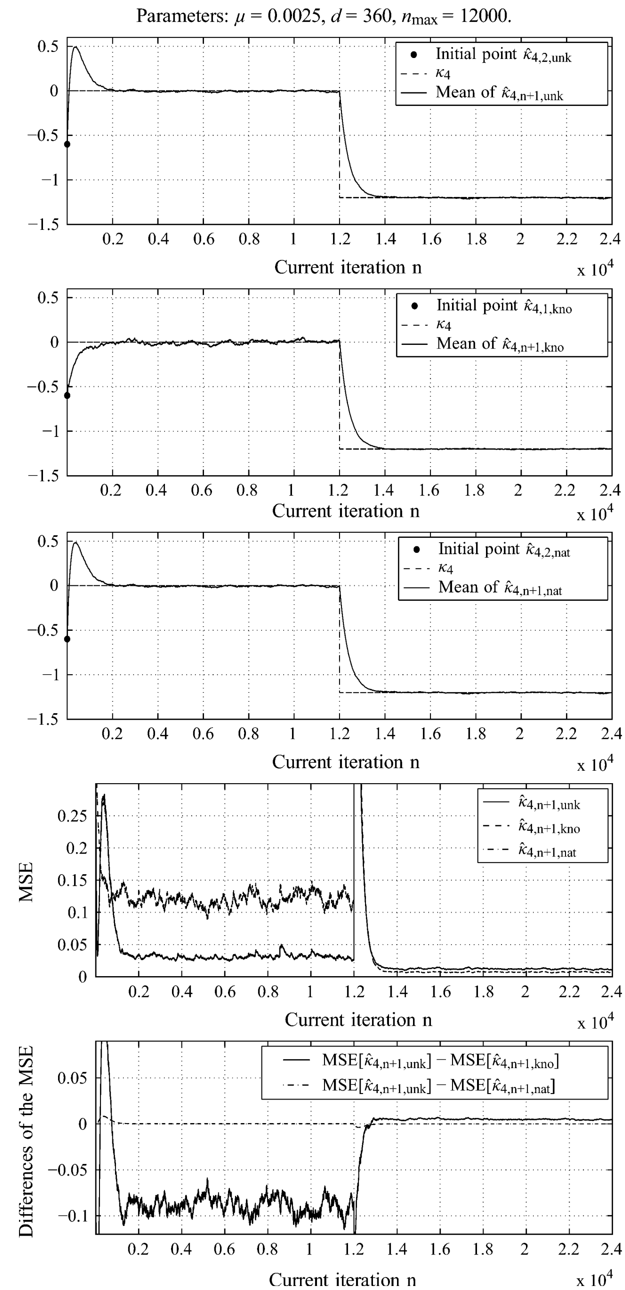


Fig. 8. Static power case: mean behavior and MSE of estimators. Note that the MSE of the statistics $\hat{\kappa}_{4,n+1,\mathrm{unk}}$ and $\hat{\kappa}_{4,n+1,\mathrm{nat}}$ are so close that they are practically undistinguishable.

[5]Formula (3) may be also written: $k_4 = \{(n^3+n^2)s_4 - 4(n^2+n)s_3 s_1 - 3(n^2-n)s_2^2 + 12 n s_2 s_1^2 - 6 s_1^4\}/\{n(n-1)(n-2)(n-3)\}$, where the power sums $s_r$ (17) are related to the sample raw moments $m'_r$ via the following relationship: $s_r = n m'_r$, $r \in \mathbb{N}^*$, [16].

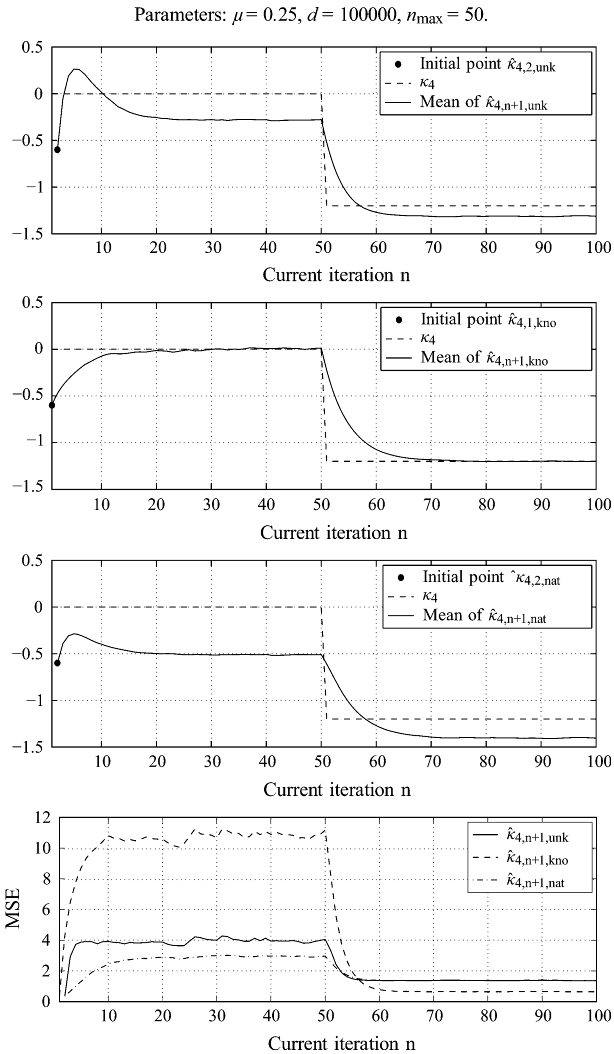


Fig. 9. Static power case: Mean behavior and MSE of estimators.

estimators may be observed in dynamic power case, which we come to study later, Figs. 10 and 11.

In order to see the consequences of a greater adaptive step size, we carried out another experience with $\mu = 0.25$, Fig. 9. The convergence is faster, it is attempted after 20 iterations, but the variance and MSE, are greater. Actually, in general, this is the common property for the adaptive estimation: the smaller the step size $\mu$, the slower the convergence and the smaller the variance, which can be the oscillations around the mean. Thus, we do not present a single realization of such statistics, but its mean and MSE. From Fig. 9, we can clearly observe the bias of the estimator for the unknown variance case and that of the natural estimator. Exactly as was predicted by the theory (see Fig. 3 and the corresponding formulas, and compare Fig. 2 with Fig. 9), the bias of the estimator for the unknown variance case is about $-0.286$ for the normal case, and $-0.114$ for the uniform case; the bias of the natural estimator is about $-0.514$ and $-0.206$, respectively, that is to say, almost two times greater in both cases. The trivial estimator $\hat{\kappa}_{4,n+1,\mathrm{kno}}$, as expected is unbiased. On the other hand, the MSE showed in Fig. 9 is slightly better for $\hat{\kappa}_{4,n+1,\mathrm{nat}}$ than for $\hat{\kappa}_{4,n+1,\mathrm{unk}}$ in the normal case, and is the same for $\hat{\kappa}_{4,n+1,\mathrm{unk}}$ and $\hat{\kappa}_{4,n+1,\mathrm{nat}}$ in the uniform case. As to the estimator $\hat{\kappa}_{4,n+1,\mathrm{kno}}$, its MSE is again strongly influenced by the parent distribution of $x_i$: it can be better (uniform case) or worse (normal case) than that of two other estimators. Thus, the proposed estimator $\hat{\kappa}_{4,n+1,\mathrm{unk}}$ showed good results in the adaptive version as well, and with respect to the natural estimator, it may be attractive especially because of its bias performances (which are twice better), while it has variable MSE performances. Note also, that our theoretical calculations represent very exactly the experimental behavior of the statistics. On the one hand, for the mean and bias, the maxima observed empirically correspond exactly to those obtained theoretically (see Fig. 2), and the asymptotical bias corresponds exactly to the observed one (see Fig. 3). On the other hand, the tedious theoretical calculations of the MSE are also in excellent accordance with experience.

We consider now the dynamic power case, Figs. 10 and 11. In order to keep the bias as low as possible, we take again $\mu = 0.0025$. The evolution of the initial power is the same as described before (33), where $k$ and $k_{\max}$ are, respectively, replaced by $n$ and $n_{\max}$. We had to choose a such big number of iterations because of the slow convergence of the algorithm for the small step sizes $\mu$. We observe that all the estimators follow quite good the changes of the current value of cumulant, provoked by the changes of the current initial input power. A little time shift between real and estimated values is due to the rate of convergence, which depends, *inter alia*, on $\mu$. Therefore, the estimators have good dynamic properties allowing them to estimate fairly accurate the fourth-order cumulant of the random signals of nonconstant power, which are obviously more frequent in practice.

*2) Second Unbiased Case $\mu = 1$:* Obviously, since for this case the variances are relatively great (see Fig. 4), we present only the means and the MSE.

First, we present the static power case, Fig. 12. Since the convergence rate is great, we took a maximum number of iterations $n_{\max} = 25$, for both $N(0,1)$ and $U(0,1)$ scenarios. As predicted, the biases are null for all three estimators. We notice the convergence of the estimator for the unknown variance case is attempted after one single iteration, while for the others, it is attempted only after 6–9 iterations. In fact, we run many simulation of such a kind, and this is not an accident: the proposed estimator converges always in one iteration, while the other estimators, in 6–9 iterations. As to the MSE of the estimators, for $\hat{\kappa}_{4,n+1,\mathrm{unk}}$ it is slightly better in $N(0,1)$ case and slightly worse in $U(0,1)$ case, compared to $\hat{\kappa}_{4,n+1,\mathrm{nat}}$ estimator. In this case, the main advantage of $\hat{\kappa}_{4,n+1,\mathrm{unk}}$ over $\hat{\kappa}_{4,n+1,\mathrm{nat}}$ is the great rate of convergence in mean and the stability of mean (it does not have a brutal transition into a false direction when the distribution law changes). As to the trivial estimator $\hat{\kappa}_{4,n+1,\mathrm{kno}}$, it showed better MSE performances than two other estimators.

Now, we present the dynamic power case, Fig. 13, with the same parameters as just before. The power evolution is again given by the amplitude modulation law (33), where $k$ and $k_{\max}$ are, respectively, replaced by $n$ and $n_{\max}$. We observe in these fast dynamic conditions, estimator $\hat{\kappa}_{4,n+1,\mathrm{unk}}$ behaves good giving a unit shift from the true cumulant value, $\hat{\kappa}_{4,n+1,\mathrm{kno}}$ behaves acceptably, while $\hat{\kappa}_{4,n+1,\mathrm{nat}}$ is completely unable to converge for the normal process, but it still converges well for

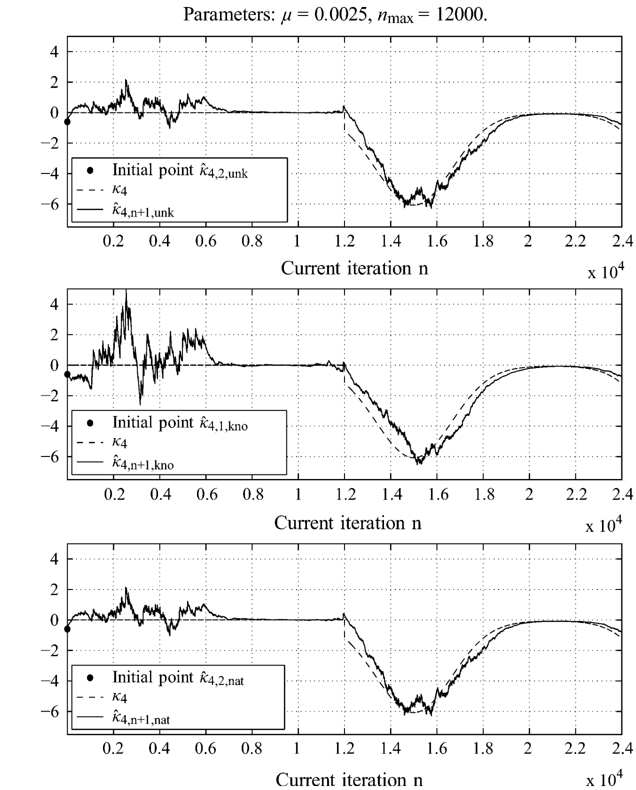


Fig. 10. Dynamic power case: single realizations of estimators.

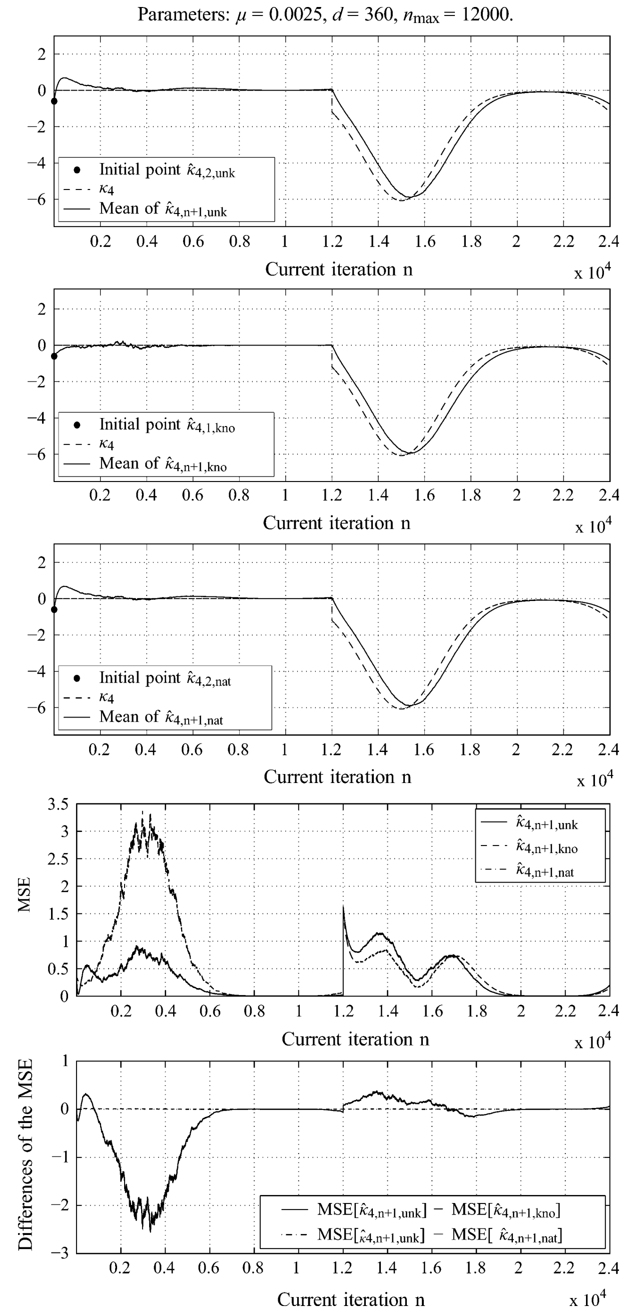


Fig. 11. Dynamic power case: mean behavior and MSE of estimators. Note that the MSE of the statistics $\hat{\kappa}_{4,n+1,\mathrm{unk}}$ and $\hat{\kappa}_{4,n+1,\mathrm{nat}}$ are so close that they are practically undistinguishable.

the uniform one. By decreasing the frequency of modulation (e.g., by taking $n_{\max} \geqslant 200$ instead of 25), the estimators $\hat{\kappa}_{4,n+1,\mathrm{kno}}$, and especially $\hat{\kappa}_{4,n+1,\mathrm{nat}}$, behave much better and their performances become quite similar to those of $\hat{\kappa}_{4,n+1,\mathrm{unk}}$. Thus, the estimator $\hat{\kappa}_{4,n+1,\mathrm{unk}}$ was found to be more robust and accurate in fast dynamic processes.

On the other hand, for this unbiased case, such MSE for all three estimators may in principle limit their practical uses. Perhaps, it could be used in some fast dynamic cyclic (or quasi-cyclic) process with identical statistical properties of each cycle (e.g., radar, lidar, or sonar signals).

Finally, we would like to note that the proposed estimator was designed according to the zero-bias principle, not to the minimum MSE one, that is why the MSE is not always optimum. On the other hand, certainly, the MSE is an important estimator's criterion, but as we could ascertain, for the known variance case, the MSE of the trivial estimator $\hat{\kappa}_{4,\mathrm{kno}}$ is much worse in normal case, while its construction is really trivial and do not casts doubts (everyone would use it, if the power was known). So, in fact, the main problem is that the MSE performances depend on the distribution laws, while the proposed estimator has always zero bias (batch version) or smaller one (adaptive versions[6]) for any law; the latter may be especially appreciated in the applications that cannot tolerate the bias, e.g., blind source separation problem [4], speech processing [8], where the performance indexes are based on the simple Monte Carlo runs (i.e., on the sample mean of cumulant). Besides, we studied only two particular distribution laws, and in order to truly judge the MSE qualities of all studies estimators, other distributions (e.g., Rayleigh, Maxwell-Boltzmann, $\chi$, different multimodal distributions, etc.) have to be considered as well.

[6]From (24) and (30), it is straightforward that for $0 < \mu < 2$, we always have $\lim_{n\to\infty}(|b_{\mathrm{nat}}| - |b_{\mathrm{unk}}|) \geqslant 0$, since $\mathrm{E}[x^4] \geqslant \mathrm{E}^2[x^2]$ (equalities when $\mu = 1, \mu \to +0$ and for any $\mu$ when the distribution law becomes a Dirac delta function or a linear combination of them, e.g., Bernoulli distribution with equiprobable states).

## V. Conclusion

We have proposed a consistent efficient estimator of the fourth-order cumulant (or semi-invariant) for real discrete-time random i.i.d. (at least up to order 8) zero-mean signal.

The first stage of our work was the elaboration of the batch version of the estimator, according to the principle of zero bias.

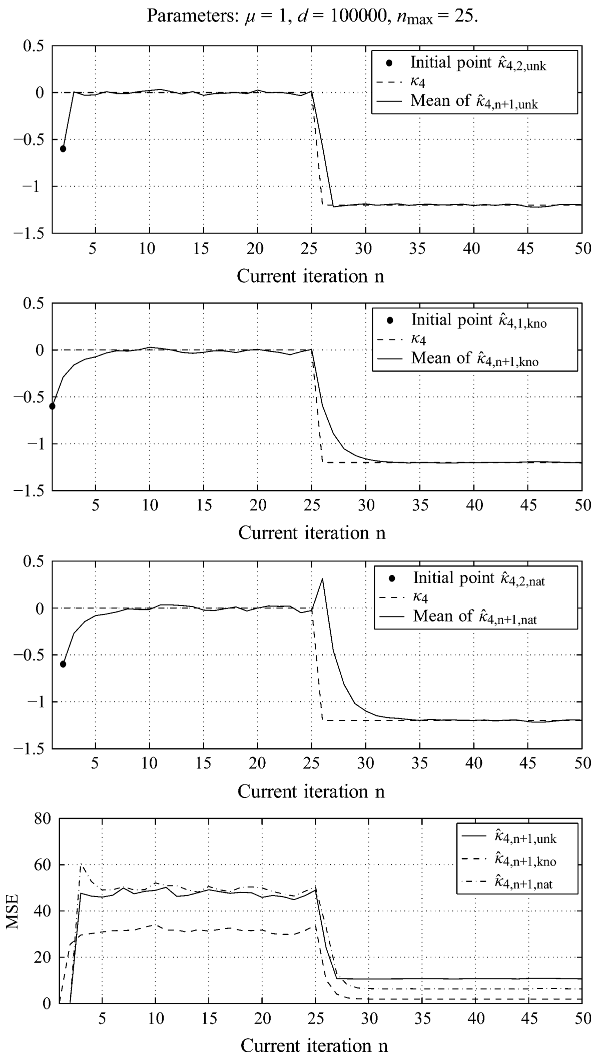


Fig. 12. Static power case: mean behavior and MSE of estimators.

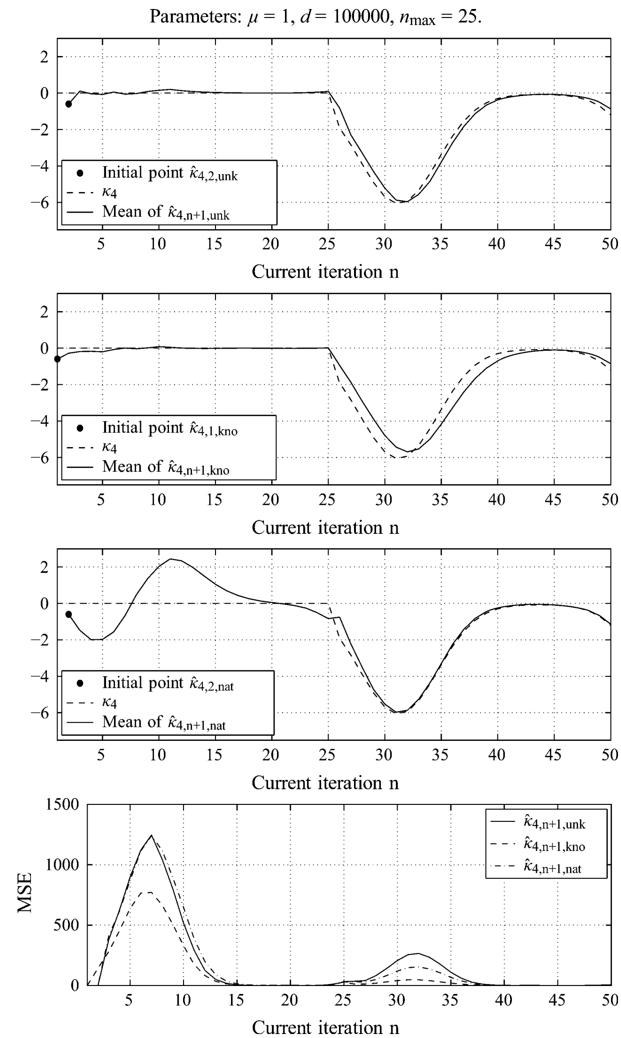


Fig. 13. Dynamic power case: mean behavior and MSE of estimators.

The elaborated estimator was found to be not only unbiased, but also quite efficient. We actually undertook a systematical theoretical and empirical studies for normal and uniform processes in static and dynamic power conditions, with comparisons to other estimation techniques of the fourth-order cumulant (natural estimator, trivial unbiased estimator for the known initial power case and the fourth $k$-statistics), and the proposed estimator showed better of equal performances than the latter ones, except the trivial estimator for the known initial power case whose performances vary a lot, depending on the distribution of the random samples $x_i$.

The second stage of our work was to provide, for all studied estimators (except the fourth $k$-statistics), their adaptive versions, and then, to study and compare them in detail. We performed for them the analysis of convergence in mean, and asymptotically, the analysis of convergence in mean square. These analyses show that the proposed estimator and the natural one are quasi-convergent in mean and in mean square, while the trivial estimator for the known variance case is convergent in mean and quasi-convergent in mean square. On the other hand, the bias of the proposed estimator is about twice smaller than that of the natural estimator for the normal and uniform distributions. Two particular unbiased cases were found: $\mu = +0$ and $\mu = 1$. Then, we performed an experimental study of the considered estimators for the normal and uniform distributions in static and dynamic power conditions. The first unbiased case was found to be the classic one, and it gives a slow rate of convergence with a small bias and small MSE. In this case, in both static and dynamic power conditions, the proposed estimator and the natural one behave practically equally, while the performances of the estimator for the known variance case depend again on the distribution of the samples $x_i$. In the intermediate case for average step sizes $\mu$, the proposed estimator was found to have its bias smaller than that of the natural one, while its efficiency may be slightly worse (normal case), or slightly better (uniform case). Again, the behavior of the estimator for the known variance case is different. Finally, the second unbiased case $\mu = 1$, gives a very fast convergence rate for the proposed estimator (one iteration is sufficient) for the normal and uniform processes in static and dynamic power conditions, while the natural estimator and the trivial one converge more slowly or are not capable at all to converge. The latter drawback becomes especially important for the natural estimator in fast dynamic power condition for the normal distribution, but it almost disappears in slow dynamic power conditions. However, all three estimators showed relatively great MSE for this second unbiased case.

## Appendix A
## Calculation of the Variance and MSE of the Batch Estimator $\hat{\kappa}_{4,\mathrm{unk}}$

The calculation of the variance and MSE is mainly reduced to the term $\mathrm{E}[\hat{\kappa}^2_{4,\mathrm{unk}}]$, because of the unbiasedness of the estimator. First, we write the term $\hat{\kappa}^2_{4,\mathrm{unk}}$ from (9) in the following form:

$$\hat{\kappa}^2_{4,\mathrm{unk}} = \alpha^2 \sum_{i=1}^{n}\sum_{j=1}^{n} x_i^4 x_j^4 - 2\alpha\beta \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} x_i^4 x_j^2 x_k^2 + \beta^2 \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} x_i^2 x_j^2 x_k^2 x_l^2$$

where the coefficients $\alpha$ and $\beta$ are given by (10). Then, we calculate the mathematical expectation of each term. By treating separately each term, and by finding into each multiple sum all crossed terms (i.e., when two or more indexes in multiple sums coincide, exactly as it was done in the (5)[7]) we can get the result. Thus, for the first term of the last expression we have

$$\begin{aligned}\mathrm{E}\Bigg[\sum_{i,j} x_i^4 x_j^4\Bigg] &= \mathrm{E}\Bigg[\sum_{i} x_i^8 + \sum_{\substack{i,j\\ i\neq j}} x_i^4 x_j^4\Bigg] \\ &= n\mathrm{E}[x^8] + n(n-1)\mathrm{E}^2[x^4].\end{aligned}$$

For the middle one

$$\begin{aligned}&\mathrm{E}\Bigg[\sum_{i,j,k} x_i^4 x_j^2 x_k^2\Bigg] \\ &\quad= \mathrm{E}\Bigg[\sum_{i} x_i^8 + 2\sum_{\substack{i,j\\ i\neq j}} x_i^6 x_j^2 + \sum_{\substack{i,j\\ i\neq j}} x_i^4 x_j^4 + \sum_{\substack{i,j,k\\ i\neq j\neq k}} x_i^4 x_j^2 x_k^2\Bigg] \\ &\quad= n\mathrm{E}[x^8] + 2n(n-1)\mathrm{E}[x^6]\mathrm{E}[x^2] + n(n-1)\mathrm{E}^2[x^4] \\ &\qquad+ n(n-1)(n-2)\mathrm{E}[x^4]\mathrm{E}^2[x^2].\end{aligned}$$

And finally, for the last term, we have the following mathematical expectation:

$$\begin{aligned}&\mathrm{E}\Bigg[\sum_{i,j,k,l} x_i^2 x_j^2 x_k^2 x_l^2\Bigg] \\ &\quad= n\mathrm{E}[x^8] + 4n(n-1)\mathrm{E}[x^6]\mathrm{E}[x^2] \\ &\qquad+ 3n(n-1)\mathrm{E}^2[x^4] + 6n(n-1)(n-2)\mathrm{E}[x^4]\mathrm{E}^2[x^2] \\ &\qquad+ n(n-1)(n-2)(n-3)\mathrm{E}^4[x^2].\end{aligned}$$

From these three latter equations, we obtain $\mathrm{E}[\hat{\kappa}^2_{4,\mathrm{unk}}]$, and consequently, (16).

[7]This method is also very well explained in [18], in the chapter related to the approximations to sampling distributions.

## Appendix B
## Calculation of the Mean of the Adaptive Estimator $\hat{\kappa}_{4,n+1,\mathrm{unk}}$

It is straightforward that since the estimator $\hat{\kappa}_{4,n+1,\mathrm{unk}}$ (20) depends on the auxiliary estimators $\hat{\sigma}^2_n$ and $\hat{\sigma}^4_n$, we should better start our calculations with these auxiliary estimators.

In the first instance, we deal with $\hat{\sigma}^2_n$. At first, we express the $n\mathrm{th}$ value of the estimator in terms of the initial value $\hat{\sigma}^2_1$, which is a deterministic one. It can be achieved by recursively decreasing the order of estimator in the right part of (22)

$$\begin{aligned}\hat{\sigma}^2_n &= (1-\mu)\hat{\sigma}^2_{n-1} + \mu x_n^2 = (1-\mu)^2\hat{\sigma}^2_{n-2} \\ &\quad+ \mu(1-\mu)x^2_{n-1} + \mu x_n^2 = \dots \\ &= (1-\mu)^{n-1}\hat{\sigma}^2_1 + \mu\sum_{k=0}^{n-2}(1-\mu)^k x^2_{n-k}.\end{aligned}$$

Then, we apply the operator of the mathematical expectation to the both sides of the last expression. On the one hand, it permits us to get rid of the random values $x_n^2$ by replacing them by the corresponding moments; on the other hand, the finite sum becomes a simple geometric series which can be easily calculated:

$$\mathrm{E}[\hat{\sigma}^2_n] = (1-\mu)^{n-1}\hat{\sigma}^2_1 + \Big[1-(1-\mu)^{n-1}\Big]\mathrm{E}[x^2]. \tag{35}$$

Finally, since we are also interested in the asymptotical mean (especially for the further calculations in the Appendix C), we calculate the limit of the last expression when $n\to\infty$

$$\lim_{n\to\infty}\mathrm{E}[\hat{\sigma}^2_n] = \begin{cases}\mathrm{E}[x^2] & \text{if } |1-\mu|<1 \\ \text{diverges} & \text{otherwise.}\end{cases} \tag{36}$$

So, the auxiliary adaptive estimator $\hat{\sigma}^2_n$ is convergent and without bias if $0<\mu<2$.

In the second instance, we deal with $\hat{\sigma}^4_n$. By proceeding in the similar way to $\hat{\sigma}^2_n$, we first write from (22)

$$\begin{aligned}\hat{\sigma}^4_n &\equiv \left(\hat{\sigma}^2_n\right)^2 = (1-\mu)^2\hat{\sigma}^4_{n-1} + 2\mu(1-\mu)\hat{\sigma}^2_{n-1}x_n^2 + \mu^2 x_n^4 \\ &= \dots = (1-\mu)^{2n-2}\hat{\sigma}^4_1 + \mu^2\sum_{k=0}^{n-2}(1-\mu)^{2k}x^4_{n-k} \\ &\quad+ 2\mu(1-\mu)\sum_{k=0}^{n-2}(1-\mu)^{2k}x^2_{n-k}\hat{\sigma}^2_{n-k-1}.\end{aligned} \tag{37}$$

Then, by calculating its mean, we obtain

$$\begin{aligned}\mathrm{E}[\hat{\sigma}^4_n] &= (1-\mu)^{2n-2}\hat{\sigma}^4_1 + \mu^2\mathrm{E}[x^4]\sum_{k=0}^{n-2}(1-\mu)^{2k} \\ &\quad+ 2\mu(1-\mu)\mathrm{E}[x^2]\sum_{k=0}^{n-2}(1-\mu)^{2k}\mathrm{E}[\hat{\sigma}^2_{n-k-1}] \\ &= (1-\mu)^{2n-2}\hat{\sigma}^4_1 + \frac{1-(1-\mu)^{2n-2}}{2-\mu} \\ &\quad\cdot\Big\{\mu\mathrm{E}[x^4] + 2(1-\mu)\mathrm{E}^2[x^2]\Big\} + 2\mathrm{E}[x^2](1-\mu)^{n-1} \\ &\quad\cdot\Big(1-(1-\mu)^{n-1}\Big)\Big(\hat{\sigma}^2_1 - \mathrm{E}[x^2]\Big)\end{aligned} \tag{38}$$

where we used (35) in order to calculate $\mathrm{E}[\hat{\sigma}^2_{n-k-1}]$. Finally, by taking the limit when $n \to \infty$ we obtain

$$\lim_{n\to\infty} \mathrm{E}[\hat{\sigma}_n^4] = \frac{\mu}{2-\mu}\mathrm{E}[x^4] + \frac{2(1-\mu)}{2-\mu}\mathrm{E}^2[x^2]. \tag{39}$$

if $0 < \mu < 2$. Note also that the case $\mu = 1$ is a particularly interesting case, because in this case the variance $\mathrm{E}[x^2]$ does not have any influence at all. Also the case $\mu \to +0$ is quite particular[8] because

$$\lim_{\substack{n\to\infty\\ \mu\to+0}} \mathrm{E}[\hat{\sigma}_n^4] = \mathrm{E}^2[x^2]. \tag{40}$$

that gives the result intuitively expected, and in this case the estimator $\hat{\sigma}_n^4$ becomes asymptotically unbiased for $\mathrm{E}^2[x^2]$. Whence, we can already suppose, that if we wish our adaptive estimators to be unbiased, $\mu$ has to be chosen small enough.

Finally, we can proceed with the adaptive estimator of cumulant itself. By using the same method, we first express the $(n+1)$th value of the adaptive estimator in terms of the initial value

$$\begin{aligned}\hat{\kappa}_{4,n+1,\mathrm{unk}} = f_0^{n-1}\hat{\kappa}_{4,2,\mathrm{unk}} &+ \frac{3\mu}{(1+\mu)(1+2\mu)}\sum_{k=0}^{n-2} f_0^k \hat{\sigma}^4_{n-k}\\ &- \frac{6\mu}{1+\mu}\sum_{k=0}^{n-2} f_0^k x^2_{n-k+1}\hat{\sigma}^2_{n-k}\\ &+ \frac{\mu}{1+\mu}\sum_{k=0}^{n-2} f_0^k x^4_{n-k+1}\end{aligned} \tag{41}$$

where we denoted

$$f_0 \equiv \frac{(1-\mu)(1+3\mu)}{(1+\mu)(1+2\mu)}. \tag{42}$$

Then, we calculate its mean

$$\begin{aligned}\mathrm{E}[\hat{\kappa}_{4,n+1,\mathrm{unk}}] = f_0^{n-1}\hat{\kappa}_{4,2,\mathrm{unk}} &+ \frac{3\mu}{(1+\mu)(1+2\mu)}\sum_{k=0}^{n-2} f_0^k \mathrm{E}[\hat{\sigma}^4_{n-k}]\\ &- \frac{6\mu}{1+\mu}\mathrm{E}[x^2]\sum_{k=0}^{n-2} f_0^k \mathrm{E}[\hat{\sigma}^2_{n-k}]\\ &+ \frac{\mu(1-f_0^{n-1})}{(1+\mu)(1-f_0)}\mathrm{E}[x^4].\end{aligned} \tag{43}$$

By substituting (35) and (38) into (43) we reach the final objective, and, thus, obtain the mean given by (23), as well as the asymptotical mean given by (24). In the analogous manner, the calculation of the means is also performed for the natural estimator (28); these results are reported in (29) and (30).

## Appendix C
## Calculation of the Asymptotical MSE of the Adaptive Estimator $\hat{\kappa}_{4,n+1,\mathrm{unk}}$

The calculation of the asymptotical MSE is mainly reduced to the asymptotical mean of the quadratic term $\hat{\kappa}^2_{4,n+1,\mathrm{unk}}$. From (20), analogously to the method employed in the Appendix B, we can write $\hat{\kappa}^2_{4,n+1,\mathrm{unk}}$ in terms of its initial value

$$\begin{aligned}&\hat{\kappa}^2_{4,n+1,\mathrm{unk}}\\ &\quad= f_0^{2n-2}\hat{\kappa}^2_{4,2,\mathrm{unk}} + \left(f_2^2+2f_1f_3\right)\sum_{k=0}^{n-2} f_0^{2k}\hat{\sigma}^4_{n-k}x^4_{n-k+1}\\ &\qquad+ f_1^2\sum_{k=0}^{n-2} f_0^{2k}\hat{\sigma}^8_{n-k} + f_3^2\sum_{k=0}^{n-2} f_0^{2k}x^8_{n-k+1}\\ &\qquad+2f_0f_1\sum_{k=0}^{n-2} f_0^{2k}\hat{\kappa}_{4,n-k,\mathrm{unk}}\hat{\sigma}^4_{n-k}\\ &\qquad+2f_1f_2\sum_{k=0}^{n-2} f_0^{2k}\hat{\sigma}^6_{n-k}x^2_{n-k+1}\\ &\qquad+2f_0f_2\sum_{k=0}^{n-2} f_0^{2k}\hat{\kappa}_{4,n-k,\mathrm{unk}}\hat{\sigma}^2_{n-k}x^2_{n-k+1} + 2f_2f_3\\ &\qquad\cdot\sum_{k=0}^{n-2} f_0^{2k}\hat{\sigma}^2_{n-k}x^6_{n-k+1} + 2f_0f_3\sum_{k=0}^{n-2} f_0^{2k}\hat{\kappa}_{4,n-k,\mathrm{unk}}x^4_{n-k+1}\end{aligned}$$

where the coefficients $f_1, f_2, f_3$ are

$$\begin{aligned}f_1 &= \frac{3\mu}{(1+\mu)(1+2\mu)}\\ f_2 &= -\frac{6\mu}{1+\mu}\\ f_3 &= \frac{\mu}{1+\mu}.\end{aligned}$$

Then, we calculate its asymptotical mean

$$\begin{aligned}&\lim_{n\to\infty}\mathrm{E}[\hat{\kappa}^2_{4,n+1,\mathrm{unk}}]\\ &\quad= \frac{1}{1-f_0^2}\Big\{\left(f_2^2+2f_1f_3\right)\mathrm{E}[x^4]\lim_{n\to\infty}\mathrm{E}[\hat{\sigma}^4_n] + f_1^2\lim_{n\to\infty}\mathrm{E}[\hat{\sigma}^8_n]\\ &\qquad+ f_3^2\mathrm{E}[x^8] + 2f_0f_1\lim_{n\to\infty}\mathrm{E}[\hat{\kappa}_{4,n,\mathrm{unk}}\hat{\sigma}^4_n]\\ &\qquad+ 2f_1f_2\mathrm{E}[x^2]\lim_{n\to\infty}\mathrm{E}[\hat{\sigma}^6_n]\\ &\qquad+ 2f_0f_2\mathrm{E}[x^2]\lim_{n\to\infty}\mathrm{E}[\hat{\kappa}_{4,n,\mathrm{unk}}\hat{\sigma}^2_n]\\ &\qquad+ 2f_2f_3\mathrm{E}[x^6]\mathrm{E}[x^2]\\ &\qquad+ 2f_0f_3\mathrm{E}[x^4]\lim_{n\to\infty}\mathrm{E}[\hat{\kappa}_{4,n,\mathrm{unk}}]\Big\}.\end{aligned} \tag{44}$$

Thus, the calculation of the asymptotical MSE of $\hat{\kappa}^2_{4,n+1,\mathrm{unk}}$ requires not only the asymptotical means of the auxiliary statistics $\hat{\sigma}_n^2$ and $\hat{\sigma}_n^4$, but also those of $\hat{\sigma}_n^6$ and $\hat{\sigma}_n^8$, as well as, those of two mixed estimators[9] $\hat{\kappa}_{4,n,\mathrm{unk}}\hat{\sigma}_n^2$ and $\hat{\kappa}_{4,n,\mathrm{unk}}\hat{\sigma}_n^4$.

First, we deal with $\hat{\sigma}_n^6$. Analogously to the method employed for the estimator $\hat{\sigma}_n^4$ in (37), we first write from (22)

$$\begin{aligned}\hat{\sigma}_n^6 &= (1-\mu)^3\hat{\sigma}^6_{n-1} + 3\mu(1-\mu)^2\hat{\sigma}^4_{n-1}x_n^2\\ &\quad+ 3\mu^2(1-\mu)\hat{\sigma}^2_{n-1}x_n^4 + \mu^3x_n^6 = \dots\end{aligned}$$

[8]The strict case $\mu = 0$ cannot be considered here, because the first limit for $n \to \infty$ will not converge for $\mu = 0$.

[9]Note that $\mathrm{E}[\hat{\kappa}_{4,n,\mathrm{unk}}\hat{\sigma}_n^2] \neq \mathrm{E}[\hat{\kappa}_{4,n,\mathrm{unk}}]\mathrm{E}[\hat{\sigma}_n^2]$ and $\mathrm{E}[\hat{\kappa}_{4,n,\mathrm{unk}}\hat{\sigma}_n^4] \neq \mathrm{E}[\hat{\kappa}_{4,n,\mathrm{unk}}]\mathrm{E}[\hat{\sigma}_n^4]$, because the estimators' indexes coincide, and thus, the estimators use the same $x_i$.

$$= (1-\mu)^{3n-3}\hat{\sigma}_1^6 + 3\mu(1-\mu)^2 \cdot \sum_{k=0}^{n-2}(1-\mu)^{3k}\hat{\sigma}_{n-k-1}^4 x_{n-k}^2 + 3\mu^2(1-\mu) \cdot \sum_{k=0}^{n-2}(1-\mu)^{3k}\hat{\sigma}_{n-k-1}^2 x_{n-k}^4 + \mu^3\sum_{k=0}^{n-2}(1-\mu)^{3k}x_{n-k}^6.$$

By calculating the asymptotical mean of the latter, provided that $0<\mu<2$ and with the help of (36) and (39), we obtain

$$\lim_{n\to\infty}\mathrm{E}[\hat{\sigma}_n^6] = \frac{1}{(2-\mu)(\mu^2-3\mu+3)} \cdot\Big\{\mu^2(2-\mu)\mathrm{E}[x^6]+3\mu(1-\mu)(3-2\mu)\mathrm{E}[x^4]\mathrm{E}[x^2] + 6(1-\mu)^3\mathrm{E}^3[x^2]\Big\}. \quad (45)$$

Thus, this statistics is biased and as previously, two interesting limit cases are present

$$\lim_{\substack{n\to\infty\\ \mu\to+0}}\mathrm{E}[\hat{\sigma}_n^6] = \mathrm{E}^3[x^2]$$

and

$$\lim_{\substack{n\to\infty\\ \mu\to1}}\mathrm{E}[\hat{\sigma}_n^6] = \mathrm{E}[x^6]$$

The former represents the asymptotic unbiasedness for the statistics $\mathrm{E}^3[x^2]$, which is again an intuitively expected result; the latter is asymptotically unbiased for $\mathrm{E}[x^6]$.

Second, we deal with $\hat{\sigma}_n^8$. Similarly to the previous lines, we write $\hat{\sigma}_n^8$ in terms of its initial value

$$\hat{\sigma}_n^8 = (1-\mu)^{4n-4}\hat{\sigma}_1^8 + 4\mu(1-\mu)^3\sum_{k=0}^{n-2}(1-\mu)^{4k}\hat{\sigma}_{n-k-1}^6 x_{n-k}^2 + 6\mu^2(1-\mu)^2\sum_{k=0}^{n-2}(1-\mu)^{4k}\hat{\sigma}_{n-k-1}^4 x_{n-k}^4 + \mu^4\sum_{k=0}^{n-2}(1-\mu)^{4k}x_{n-k}^8 + 4\mu^3(1-\mu)\sum_{k=0}^{n-2}(1-\mu)^{4k}\hat{\sigma}_{n-k-1}^2 x_{n-k}^6$$

and then, we calculate its asymptotical mean value

$$\lim_{n\to\infty}\mathrm{E}[\hat{\sigma}_n^8] = \frac{1}{(2-\mu)(\mu^2-2\mu+2)} \Big\{\mu^3\mathrm{E}[x^8] + \frac{4\mu^2(1-\mu)(2\mu^2-5\mu+4)}{\mu^2-3\mu+3}\mathrm{E}[x^6]\mathrm{E}[x^2] + \frac{6\mu^2(1-\mu)^2}{2-\mu}\mathrm{E}^2[x^4] + \frac{12\mu(1-\mu)^3(3\mu^2-8\mu+6)}{(2-\mu)(\mu^2-3\mu+3)}\mathrm{E}[x^4]\mathrm{E}^2[x^2] + \frac{24(1-\mu)^6}{(2-\mu)(\mu^2-3\mu+3)}\mathrm{E}^4[x^2]\Big\} \quad (46)$$

provided that $0<\mu<2$. The corresponding double limits are, respectively

$$\lim_{\substack{n\to\infty\\ \mu\to+0}}\mathrm{E}[\hat{\sigma}_n^8] = \mathrm{E}^4[x^2]$$

and

$$\lim_{\substack{n\to\infty\\ \mu\to1}}\mathrm{E}[\hat{\sigma}_n^8] = \mathrm{E}[x^8].$$

Third, we deal with $\hat{\kappa}_{4,n,\mathrm{unk}}\hat{\sigma}_n^2$. The previous method gives the following asymptotical mean:

$$\lim_{n\to\infty}\mathrm{E}[\hat{\kappa}_{4,n,\mathrm{unk}}\hat{\sigma}_n^2] = \frac{1}{1-(1-\mu)f_0} \cdot\Big\{(1-\mu)f_1\lim_{n\to\infty}\mathrm{E}[\hat{\sigma}_n^6] + \mu f_0\mathrm{E}[x^2]\lim_{n\to\infty}\mathrm{E}[\hat{\kappa}_{4,n,\mathrm{unk}}] + \big(\mu f_1+(1-\mu)f_2\big)\mathrm{E}[x^2]\lim_{n\to\infty}\mathrm{E}[\hat{\sigma}_n^4] + \big(\mu f_2+(1-\mu)f_3\big)\mathrm{E}[x^4]\mathrm{E}[x^2]+\mu f_3\mathrm{E}[x^6]\Big\} \quad (47)$$

where the corresponding limits are given according to the above calculated asymptotical means (23), (39), and (45).

Fourthly, we deal with $\hat{\kappa}_{4,n,\mathrm{unk}}\hat{\sigma}_n^4$. By using the same method, its asymptotical mathematical expectation yields

$$\lim_{n\to\infty}\mathrm{E}[\hat{\kappa}_{4,n,\mathrm{unk}}\hat{\sigma}_n^4] = \frac{1}{1-(1-\mu)^2 f_0} \cdot\Big\{(1-\mu)^2 f_1\lim_{n\to\infty}\mathrm{E}[\hat{\sigma}_n^8] + \mu^2 f_0\mathrm{E}[x^4]\lim_{n\to\infty}\mathrm{E}[\hat{\kappa}_{4,n,\mathrm{unk}}] + (1-\mu)\big(2\mu f_1+(1-\mu)f_2\big)\mathrm{E}[x^2]\lim_{n\to\infty}\mathrm{E}[\hat{\sigma}_n^6] + 2\mu(1-\mu)f_0\mathrm{E}[x^2]\lim_{n\to\infty}\mathrm{E}[\hat{\kappa}_{4,n,\mathrm{unk}}\hat{\sigma}_n^2] + \big(\mu^2 f_1+2\mu(1-\mu)f_2+(1-\mu)^2 f_3\big)\mathrm{E}[x^4]\lim_{n\to\infty}\mathrm{E}[\hat{\sigma}_n^4] + \mu\big(\mu f_2+2(1-\mu)f_3\big)\mathrm{E}[x^6]\mathrm{E}[x^2]+\mu^2 f_3\mathrm{E}[x^8]\Big\} \quad (48)$$

where the corresponding limits are given according to the above-calculated asymptotical means (23), (39), (45), (46), and (47).

Thus, by substituting (23), (39), (45), (46), (47) and (48) into (44), we obtain the mathematical expectation of the quadratic term $\hat{\kappa}_{4,n+1,\mathrm{unk}}^2$. After that, the MSE is calculated according to (13) and (23). Similarly, by replacing the corresponding coefficients $f_0,\ldots,f_3$, the MSE of the natural estimator is calculated.

Last, note that the MSE analysis did not lead to the new domains of convergence, because it is always restricted by the smaller one defined by the auxiliary estimator $\hat{\sigma}_n^2$, and the analogical reasoning can be applied to the estimator $\hat{\kappa}_{4,n+1,\mathrm{nat}}$.

## REFERENCES

[1] J. K. Tugnait, "Identification and deconvolution of multichannel linear non-Gaussian processes using higher order statistics and inverse filter criteria," *IEEE Trans. Signal Process.*, vol. 45, pp. 658–672, Mar. 1997.

[2] C. Simon, Ph. Loubaton, and C. Jutten, "Separation of a class of convolutive mixtures: A contrast function approach," *Signal Process.*, vol. 81, no. 4, pp. 883–887, Apr. 2001.

[3] E. Moreau and O. Macchi, "High order contrasts for self-adaptive source separation," *Int. J. Adaptive Control Signal Process.*, vol. 10, no. 1, pp. 19–46, Jan. 1996.

[4] M. Castella, S. Rhioui, E. Moreau, and J.-C. Pesquet, "Quadratic higher order criteria for iterative blind separation of a MIMO convolutive mixture of sources," *IEEE Trans. Signal Process.*, vol. 55, pp. 218–232, 2007.

[5] P. Comon, "MA identification using fourth order cumulants," *Signal Process.*, vol. 26, no. 3, pp. 381–388, Mar. 1992.

[6] D. Dembélé and G. Favier, "Recursive estimation of the fourth-order cumulants with application to identification," *Signal Process.*, vol. 68, no. 2, pp. 127–139, 1998.

[7] K. M. Indrebo, R. J. Povinelli, and M. T. Johnson, "Third-order moments of filtered speech signals for robust speech recognition," in *Nonlinear Analyses and Algorithms for Speech Processing (Lecture Notes in Computer Science)*. Berlin/Heidelberg: Springer, 2005, pp. 151–157.

[8] L.-R. Shen, X.-Y. Li, W. Wei, R.-B. Zhang, and H.-Q. Wang, "Speech stream detection based on higher-order statistics," in *Proc. IEEE Int. Conf. Mach. Learn. Cybern.*, 2003, vol. 5, pp. 3086–3089.

[9] A. Alkulaibi, J. J. Soraghan, and T. S. Durrani, "Fast HOS-based simultaneous voiced/unvoiced detection and pitch estimation using 3-level binary speech signals," in *Proc. 8th IEEE Signal Process. Workshop on Statist. Signal Array Process. (SSAP '96)*, 1996, pp. 194–197.

[10] K. Forest, G. Weismer, P. Milencovic, and R. N. Dougall, "Statistical analysis of word-initial voiceless obstruents: Preliminary data," *J. Acoust. Soc. Amer.*, vol. 84, pp. 115–123, 1988.

[11] A. Moreno, S. Tortola, J. Vidal, and J. Fonollosa, "New HOS-based parameter estimation methods for speech recognition in noisy environnements," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 1995, pp. 429–432.

[12] M. Rangoussi, A. Delopoulos, and M. Tsatsanis, "On the use of the higher-order statistics for robust endpoint detection of speech," in *Proc. IEEE Signal Process. Workshop on Higher-Order Statist.*, 1993, pp. 56–60.

[13] J. W. Fackrell and S. McLaughlin, "The higher-order statistics of speech signals," *IEEE Colloq. Tech. Speech Process. Their Appl.*, pp. 7/1–7/6, 1994.

[14] R. A. Fisher, "Moments and product moments of sampling distributions," *Proc. London Math. Soc., Series 2*, vol. 30, pp. 199–238, 1929.

[15] E. A. Cornish and R. A. Fisher, *Moments and Cumulants in the Specification of Distributions*. Paris: Revue de l'Institut International de Statistique, 1937, vol. 5, pp. 307–320.

[16] H. Jeffreys, *Theory of Probability. International Series of Monographs on Physics*, R. H. Flower and P. Kapitza, Eds. Oxford, U.K.: Clarendon, 1939.

[17] A. N. Shiryaev, *Probability*, 2nd ed. Berlin: Springer-Verlag, 1995.

[18] M. G. Kendall, *The Advanced Theory of Statistics Vol. I. Distribution Theory*, 2nd ed. London, U.K.: Charles Griffin, 1945.

[19] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications," *Proc. IEEE*, vol. 79, pp. 278–305, 1991.

[20] P. Amblard and J.-M. Brossier, "Adaptive estimation of the fourth-order cumulant of a white stochastic process," *Signal Process.*, vol. 42, no. 1, pp. 37–43, 1995.

[21] M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics Vol. II. Inference and Relationship*, 2nd ed. London, U.K.: Charles Griffin, 1967.

[22] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991.

[23] G. Saporta, *Probabilités, Analyse des Données et Statistiques*. Paris: Éditions Technip, 1990.

[24] J. F. Kenney and E. S. Keeping, *Mathematics of Statistics. Part I*, 3rd ed. New York: Van Nostrand , 1961.

[25] J. F. Kenney and E. S. Keeping, *Mathematics of Statistics. Part II*, 2nd ed. New York: Van Nostrand, 1959.

**Iaroslav V. Blagouchine** was born in Saint-Petersburg, Russia, on December 22, 1979. He received the B.S. degree in physics from the Saint-Petersburg State University in 2000, and the M.S. degree in electronic engineering from the Grenoble Institute of Technology, France, in 2001.

From 2001 to 2002, he was with the Department Señales, Sistemas y Radiocomunicaciones, Universidad Politécnica de Madrid, Spain, and in 2003, he worked as an engineer with the Grenoble Institute of Technology. From 2004 to 2007, he was pursuing the Ph.D. degree as well as working as a Teacher Assistant in the same institute. Since 2007, he holds a Postdoctoral and Teacher Assistant positions with the Telecommunication Department, University of Toulon, France. His main research interests are in statistical signal processing, higher-order and inferential statistics, as well as the constraint optimization problems.

**Eric Moreau** (M'96–SM'08) was born in Lille, France. He graduated from the Ecole Nationale Supérieure des Arts et Métiers, Paris, France, in 1989. He received the Agrégation de Physique degree from the Ecole Normale Supérieure de Cachan, France, in 1990 and the DEA and Ph.D. degrees in signal processing from the Université Paris-Sud, France, in 1991 and 1995, respectively.

From 1995 to 2001, he was an Assistant Professor with the Telecommunications Department, Institut des Sciences de l'Ingénieur de Toulon et du Var, La Valette, France. He is currently a Professor with the University of Toulon, France. His main research interests are in statistical signal processing using higher-order statistics.