Improvement of Cumulant-based Parameter estimation

Denis Kouamé, Jean-Marc Girault

LUSSI, GIP Ultrasons Tours, France

Abstract— **This paper presents a improvement of high order cumulant-based parameter estimation using delta operator applied to instrumental variable algorithm. It is based on a modification of the classical least squares estimation and the utilization of the delta operator and the introduction of an additional term in the parameter estimates. Computer simulation results are given to illustrate the behavior of this method.**

I. INTRODUCTION

SECOND order statistics (SOS) parameter estimation, using of a δ-operator has shown advantages over the classical q-operator, in terms of numerical ability and ill-conditions processing. Moreover, it has been shown that the least squares estimation convergence rate can be improved [1]- [3] using a δ-operator . Concerning high order statistics (HOS), the major problems are the convergence rate and the fluctuations of parameter estimation when the algorithms, particularly the Recursive Instrumental Variable (RIV), work on short-length data, [4]-[6]. To this day, no δ-operator type algorithm applicable to HOS is available. Such an algorithm, by using the properties of δ-operator, could lead to an improvement of classical algorithms. The aim of this paper is to investi-gate a δ-operator high order statistics recursive instrumental variable method.

II. PRELIMINARIES

The following notations will be used. x denotes a vector, $x(n)$ denotes the nth (time) element of x. Uppercase (X) denotes a matrix. The finite difference δ-operator is defined by: $\delta x(n) = \frac{x(n)-x(n-1)}{T} = \frac{1-q}{T} x(n)$, where T is the (normalized) sampling period of the process x. q is the delay operator that is, for an integer i, $q^i x(n) = x(n-i)$. Thus, we define $\delta^i x(n) = \frac{(1-q)^i}{T^i} x(n) = \frac{1+...+(-1)^i q^i}{T^i} x(n)$. m-order cumulant of x is denoted by $C_{mx}(t_1, ...t_{m-1})$. Consider a real discrete time process,

$$x(n) = \sum_{k=1}^{p} a(k)x(n-k) + u(n), \; y(n) = x(n) + w(n) \quad (1)$$

where $u(n)$ is an independent input excitation, $y(n)$ the observed output, and $w(n)$ a zero-mean gaussian colored noise (which can be an MA process). Here, we consider the problem of estimating the AR parameter, $a(k)$, $k = 1, ..., p$, using recursive instrumental variable transversal structure. This problem has been thoroughly covered in many papers, *e.g.*[7]. The Instrumental Variable (IV) estimation of the AR part of (1) uses a process $z(n)$ referred to as instru-

mental variable, which will be discussed below. We define

$$\phi^T(n) = [y(n-1), ...y(n-p)], \psi^T(n) = [z(n-1), ...z(n-p)]$$

and $\theta^T = [a(1)...a(p)]$. The process $z(n)$, assumed to be uncorrelated to the additive noise $w(n)$, asymptotically yields the $(p+1)$ rank matrix

$$P^{-1}(n) = \sum_{k=0}^{n} \lambda^{n-k} \psi(k) \phi^T(k) \qquad (2)$$

Equation (2) gives the classical correlation-based estimate for $z(n) = y(n)$. In this case, the parameter estimates can be obtained via the Recursive Least Squares (RLS) method. To overcome RLS limitations, classically $z(n)$ is chosen so that $P(n)$ is an m-order cumulant matrix ($m > 2$). Typical choices $e.g.$ [4] are:

- $z(n) = y(n)y(n + n_0)$; this leads to estimates based on 1D slice $C_{3y}(k, k + n_0)$. Generally n_0 is set to zero and $z(n) = y^2(n)$ and then estimates are based on the diagonal slice of the 3rd-order cumulants of y.

- $z(n) = y^3(n) - 3\sigma^2 y(n)$, with $\sigma^2 = C_{2y}(0)$; this leads elements of P to be sample estimates of $C_{4y}(k, k, k)$ the diagonal slice of the 4th-oder cumulant of y.

Here, we consider the case of estimation using m-order cumulant ($m > 2$). The classical q-operator Recursive Instrumental Variable algorithm is then given by :

$$\hat{\theta}(n) = \hat{\theta}(n-1) + P(n)\psi(n)[y(n) - \phi^T(n)\theta(n-1)] \quad (3)$$

$$P(n) = \lambda^{-1}[P(n-1) - \frac{P(n-1)\psi(n)\phi^T(n)P(n-1)}{\lambda + \phi^T(n)P(n-1)\psi(n)}] \quad (4)$$

with $0 < \lambda \leq 1$ This algorithm suffers from high fluctuations related to the use of High Order Cumulants (HOC) when applied to short-length data.

III. Proposed Algorithm

Now consider the extension of model (1) using the δ-operator, which can be written as:

$$A(\delta)x(n) = u(n); \ with \ A(\delta) = \sum_{i=0}^{p} \alpha(i)\delta^i \qquad (5)$$

Define:

$$\bar{\theta}^T = [-\alpha(p-1)... - \alpha(0)] \qquad (6)$$

$$\bar{\phi}^T(n) = [\delta^{(p-1)}x(n)...x(n)] \qquad (7)$$

$$\bar{y}(n) = \delta^{(p)}x(n) + w(n) \qquad (8)$$

$$\bar{\psi}^T(n) = [\delta^{(p-1)}z(n)...z(n)] \qquad (9)$$

$Remark\ 1$: Relations between general model (5) (say $\alpha(n)$) and (1),(say $a(n)$) come immediately by expanding (5), for example

$$\hat{\theta}(1) = 2 + T\hat{\bar{\theta}}(1); \hat{\theta}(2) = -1 - T\hat{\bar{\theta}}(1) + T^2\hat{\bar{\theta}}(2)$$

It is well known that the methodology in developing IV estimates is based on modification of least squares covariance matrix so that the observations are not correlated to the noise, through introduction of an instrumental process [7]. Following this methodology, the δ-operator RIV estimates are dervived. Using $\tilde{\hat{\bar{\theta}}}(n) = \hat{\bar{\theta}}(n) - \hat{\bar{\theta}}(n-1)$ and initial cumulant matrix P_0, the orthogonality condition [7] leads to :

$$\left[P_0^{-1} + \sum_{k=1}^{n} f_n(k)\bar{\psi}(k)\bar{\phi}^T(k)\right]\tilde{\hat{\bar{\theta}}}(n) =$$

$$\sum_{k=1}^{n} f_n(k)\bar{\psi}(k)[\bar{y}(k) - \bar{\phi}^T(k)\hat{\bar{\theta}}(n-1)]$$

$$where \ f_n(k) = (1 - \bar{\lambda}T)^{n-k} \qquad (10)$$

$f_n(k)$ uses the forgetting factor $\bar{\lambda}$ to account for slow variations of the parameters. Parameter estimates are then given by :

$$\tilde{\hat{\theta}}(n) = \bar{P}(n) \sum_{k=0}^{n} f_n(k)\bar{\psi}(k) \left[\bar{y}(k) - \bar{\phi}^T(k)\tilde{\hat{\theta}}(n-1) \right]$$

$$\bar{P}(n) = \left[P_0^{-1} + \sum_{k=1}^{n} f_n(k)\bar{\psi}(k)\bar{\phi}^T(k) \right]^{-1} \quad (11)$$

Unlike the classical algorithm, here, the case without forgetting factor corresponds to $\bar{\lambda} = 0$.

For convenience, in the following parts we assume, without lost of generality, that $T = 1$. From these definitions, using similar derivations as in [8], the δ-operator RIV estimates are obtained by:

$$\delta\tilde{\hat{\theta}}(n) = \bar{P}(n)\{P_0^{-1}\delta\tilde{\hat{\theta}}(n-1) + \bar{\psi}(n)[\bar{y}(n) - \bar{\phi}^T(n)\hat{\theta}(n-1)]\}$$

$$(12)$$

$$\delta\bar{P}(n) = \frac{1}{1-\bar{\lambda}}\{\bar{\lambda}\bar{P}(n) - \frac{\bar{P}(n)\bar{\psi}(n)\bar{\phi}^T(n)\bar{P}(n)}{1+\bar{\phi}^T(n)\bar{P}(n)\bar{\psi}(n)}\} \quad (13)$$

In practice, P_0 initial cumulant marix is chosen by setting $P_0 = \gamma I$, where γ is a scalar and I the identity matrix. Here, one can notice that equations (12) and (13), by using an additional term $\bar{P}(n)P_0^{-1}\delta\hat{\theta}(n-1)$ and by estimating the finite difference of parameters and cumulant matrix, differ from equations (3) and (4). This algorithm, first results in the reduction of parameter estimates fluctuations and thus accelerates convergence speed. Secondly, it takes advantage of good numerical properties of the δ-operator [1],[3]. The algorithm can be summarized by :

• choose initial values of $\bar{P}$ (e.g. $P_0 = \gamma I$ with I the identity matrix and γ a positive constant) and $\tilde{\hat{\theta}}(0)$

For n=1,2,...

• compute $\bar{\theta}(n)$; $\bar{\phi}(n)$; $\bar{y}(n)$; $\bar{\psi}^T(n)$ eqs. (6)-(9);

• compute $\delta\bar{P}(n)$ eq. (13), and $\bar{P}(n)$ using the definition of $\delta\bar{P}(n)$ as specified in section II.

• compute $\delta\tilde{\hat{\theta}}(n)$ eq. (12) and $\tilde{\hat{\theta}}(n)$ using the definition of $\delta\bar{\theta}(n)$ as specified in section II.

• the algorithm ends here if the model used is (5);

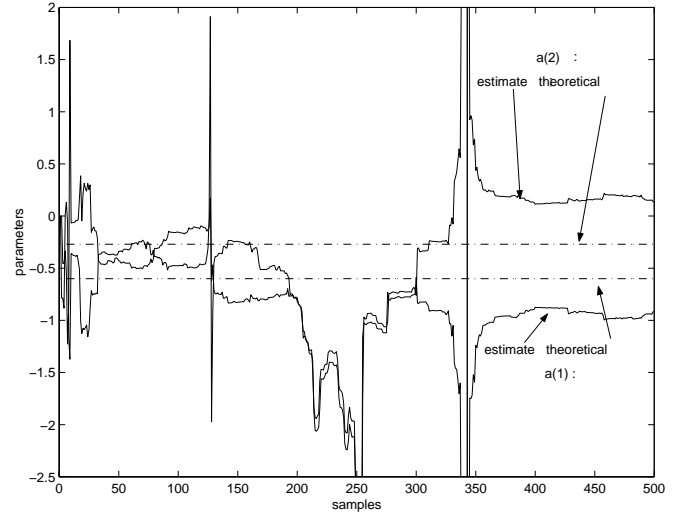• if not compute $\hat{\theta}$ from $\tilde{\hat{\theta}}$ using $Remark 1$.

Fig. 1.

Parameter estimate with classical RIV algorithm. Theoretical values (dash dot) and estimated values (solid lines) with a single realization

The counterpart of this algorithm is the computation of the additional term and finite differences. But this is a small price to pay compared to the improvement in terms of fluctuations reduction. To illustrate this, we consider for the purpose of a computer simulation the process : $x(n) - 0.6x(n-1) - 0.27x(n-2) = u(n)$; with $w(n) = -0.49w(n-2) + e(n)$; $y(n) = x(n) + w(n)$ where
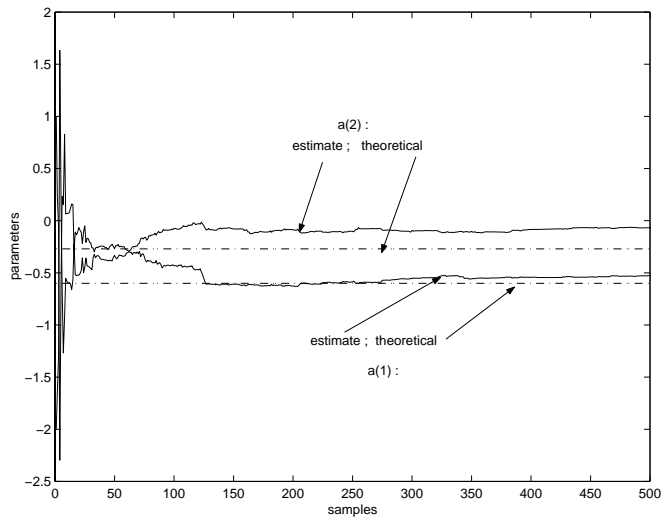
Fig. 2.

Parameter estimate with new RIV algorithm. Theoretical values (dash dot) and estimated values (solid lines) with a single realization

$u(n)$ is 2-squewness, unit variance, zero-mean exponentially distributed random process and e a gaussian noise. In figs.(1) and (2), we show the results of parameter estimates using the 3rd-order cumulant with classical RIV and the new algorithm for SNR=10 dB. For this simulation, we use $T = 1$ (chosen arbitrarily). The initial cumulant matrix for the two algorithms is chosen as $P_0 = 10^4 I$ where I is the identity matrix. $\lambda = 1$ for the classical algorithm and $\bar{\lambda} = 0$ for the new algorithm (these two choices for λ are similar). Only one realization of the signals is used for this simulation. The number of samples is 500 (whereas usually, classical algorithm necessitates higher numbers and many realizations and an average of the estimates to reduce the fluctuations). As it can be seen, even in these conditions, the new algorithm yields less fluctuations (fig.(2)), compared to classical algorithm (fig.(1)), where it is difficult to distinguish between the two parameter estimates because of the high fluctuations.

IV. Conclusion

In this paper we have presented a new parameter estimation algorithm which has been applied to HOC . This algorithm is based on the modification of the classical approach by i using a δ-operator. It allows reduction of parameter estimates fluctuations, even when short-length data is used as shown by computer simulation.

References

[1] R.H. Middleton and G.C. Goodwin, *Digital Control and Estimation. A unified approach*, Prentice Hall New Jersey, 1990.

[2] H. Fan and X. Liu, "Delta levinson and schur type rls algorithms for adaptive signal processing," *IEEE Trans. Signal Processing*, vol. 42, no. 7, pp. 1629–1639, 1994.

[3] L. Qiang, H. Fan, and E. Karlsson, "A delta mywe algorithm for parameter estimation of noisy ar process," *IEEE Trans. Signal Processing*, vol. 44, no. 5, pp. 1300–1303, 1996.

[4] D. Aboutajdine, A. Adid, and A. Meziane, "Fast adptive algorithms for ar parameter estimation using high order statistics," *IEEE Trans. Signal Processing*, vol. 44, no. 8, pp. 1998–2009, 1996.

[5] B. Friedlander and B. Porat, "Adaptive iir algorithm based on high order statistics," *IEEE Trans. Acoust. Speech and Signal Processing*, vol. 37, no. 4, pp. 485–495, 1989.

[6] X.D. Zhang, Y. Song, and Y.D. Li, "Adaptive identification of nonminimum phase models using high order cumulants alone," *IEEE Trans. Signal Processing*, vol. 44, no. 5, pp. 1285–1288, 1996.

[7] T. Soderstrom and P. Stoica, "Comparison of some instrumental variable methods - consistency and accuracy aspects," *Automatica*, vol. 17, pp. 101–115, 1981.

[8] D. Kouamé, J.F. Roux, and A. Ouahabi, "Parametric estimation: Improvement of the rls approach using a differential approach," *Int. J. of Modelling and Simulation*, vol. 19, pp. 18–26, 1999.