

NOVEL BLIND VARIANTS OF THE OBE ALGORITHM

M. Nayeri

T.M. Lin

J. R. Deller

Michigan State University
Department of Electrical Engineering / 260 EB
Speech Processing and Adaptive Signal Processing Labs
East Lansing, MI 48824-1226 USA

ABSTRACT

Set-membership algorithms, including the conventional *optimal bounding ellipsoid* (OBE) algorithm, require *a priori* knowledge of *exact* error bounds which is unknown in most applications. Conservative (over-estimated) error bounds used in practice lead to inconsistent parameter estimation. The novel OBE algorithm with automatic bound estimation (OBE-ABE) is shown to be consistently convergent without *a priori* knowledge of error bounds, even in correlated-error environments. Computationally efficient variants of this algorithm for both time-invariant and time-varying systems are presented. Simulations are performed to demonstrate the merit of the algorithms.

1. INTRODUCTION

Systems arising in signal processing, system identification, control, and communication problems can often be modeled by a *linear-in-parameters* model

$$y_n = \theta_*^T x_n + v_n \quad (1)$$

in which $\theta_* = [a_1, \dots, a_p, b_0, b_1, \dots, b_q]^T$ is the unknown "true" parameter vector to be identified; $\{x_n\}$ is a sequence of observable vectors of dimension $m = p+q+1$; and $\{v_n\}$ is an unobservable model-error sequence. An important special case is the *auto-regressive with exogenous input* (ARX) model in which $x_n = [y_{n-1}, \dots, y_{n-p}, u_n, \dots, u_{n-q}]^T$ is the observed data set composed of samples of the observable input sequence $\{u_n\}$ and output sequence $\{y_n\}$. In this paper, the ARX model is assumed and v_n is bounded with *unknown* least upper bound $\sqrt{\gamma_*}$ or greatest lower bound $-\sqrt{\gamma_*}$ for all n . The results of this paper are generalizable to cases in which y_n and v_n are vectors [2].

Conventional OBE algorithms [1, 2, 4, 5, 10] are based on the premise that v_n has a pointwise bound that is known *a priori*, $v_n^2 \leq \gamma_n$, for all n . With knowledge of energy bounds $v_{n*}^2 \leq \gamma_n$, it can be demonstrated that $\theta_* \in \Omega_n$ where Ω_n is a hyperellipsoidal set based on the observations at time n , $\Omega_n \stackrel{\text{def}}{=} \{\theta | (\theta - \theta_n)^T P_n^{-1} (\theta - \theta_n) < \kappa_n\}$ where the ellipsoid center, θ_n , and the defining matrix P_n are computed recursively. The ellipsoid center θ_n is used as an estimator of the parameters θ_* at each n . For details see, for example, [4, 2, 5]. However, since $\{v_n\}$ is unobservable, choosing a proper bounding sequence $\{\gamma_n\}$ is critical in practice. If one or more bounds are underestimated, i.e., $v_n^2 < \gamma_n$ at one or more n , then the algorithm is no longer theoretically

valid. Simulations show that underestimation results in an *inconsistent* estimator.

A conservative sequence of bounds $\{\gamma_n\}$ will assure a meaningful ellipsoid at each n . However, recent work of Nayeri *et al.* [7, 10] has demonstrated that the estimator may be very imprecise, even asymptotically, if the bounds are too "loose." Formally, it has been shown that even if a *persistence of excitation* (PE) condition holds, i.e., even if there exists an $N_1 \in \mathbb{N}$ and $\rho_1, \rho_2 > 0$ such that for all n , $0 < \rho_1 I \leq \sum_{k=n+1}^{n+N_1} x_k x_k^T \leq \rho_2 I < \infty$, the sequence of ellipsoids of SM-SA (or any OBE) does not asymptotically converge to a point if a conservative bound sequence is used. A formal statement is as follows:

Lemma 1 (Proof: [7, 10]). *Assume that PE holds. If there exists an $\epsilon > 0$ and $N \in \mathbb{N}$, such that $\gamma_n - v_n^2 > \epsilon, \forall n > N$, then the sequence of ellipsoids of OBE algorithms does not asymptotically shrink to a point.*

Lemma 1 can be extended to the following by considering an almost sure (a.s.) ω -set in the proof:

Lemma 2 *Assume that PE holds. If there exists an $\epsilon > 0$ and $N \in \mathbb{N}$, such that $\gamma_n - v_n^2 > \epsilon, \forall n > N$, almost surely, then the sequence of ellipsoids of OBE algorithms does not asymptotically shrink to a point a.s.*

Even if the true bound γ_* is known (i.e., let $\gamma_n = \gamma_*$) and if PE holds, OBE algorithms are only shown in [1] to asymptotically "converge" to a non-infinitesimal region around the true parameter vector. In fact, the consistency in probability (p. consistency) of the estimator (reduction of ellipsoids to a point in a deterministic analogy) of OBE algorithms requires one more (necessary) condition on the error sequence $\{v_n\}$ (Theorem 1 below). To the best of our knowledge, the first proofs of the p. consistency for OBE algorithms in the i.i.d. case are found in [7, 10], for the *exact parameter bounding* (EPA) algorithm in [12], for OBE-ABE algorithm in i.i.d. case in [6, 9], and for OBE-ABE algorithm in correlated-error case in this paper.

For analysis, v_n , u_n , and y_n are modeled as random variables defined on a probability space $(\Omega, \mathcal{F}, P)$ where Ω is a sample space, $\mathcal{F}$ a σ -field, and P a probability measure. For convenience, define $\mathcal{F}_n \stackrel{\text{def}}{=} \sigma\{v_m, u_{m+1}, m \leq n\}$, $D_\epsilon^+ = [\sqrt{\gamma_*} - \epsilon, \sqrt{\gamma_*}]$ and $D_\epsilon^- = [-\sqrt{\gamma_*}, -\sqrt{\gamma_*} + \epsilon]$.

Definition 1 *A random sequence $\{v_n\}$ is called uniformly conditionally tailed (UCT) if given $\epsilon > 0$, there exist a $\delta > 0$ and an infinite subsequence $\{m_i\} \subset \mathbb{N}$, such that for all $n \in \{m_i\}$, $P(v_n \in (D_\epsilon^+ \cup D_\epsilon^-) | \mathcal{F}_{n-1}) > \delta$ a.s.*

Definition 2 A random sequence $\{v_n\}$ is called uniformly tailed (UT) if given $\epsilon > 0$, there exist a $\delta > 0$ and an infinite subsequence $\{m_i\} \subset \mathbb{N}$, such that for all $n \in \{m_i\}$, $P(v_n \in (D_\epsilon^+ \cup D_\epsilon^-)) > \delta$ a.s.

We conclude this section by introducing the following theorem giving a necessary condition for the shrinking of the ellipsoids of OBE algorithms to a point a.s.

Theorem 1 Assume that PE holds. Then, UCT is a necessary condition for the sequence of the ellipsoids of any OBE algorithm to shrink to a point a.s.

Proof: Suppose that UCT does not hold. That is, there exist an $\epsilon > 0$ and an $N \in \mathbb{N}$ such that, for all $n > N$, $P(v_n \in D_\epsilon^+ | \mathcal{F}_{n-1}) = 0$ and $P(v_n \in D_\epsilon^- | \mathcal{F}_{n-1}) = 0$ a.s. Hence, $\gamma_n - v_n^2 > \epsilon$, $\forall n > N$ a.s. Thus, by Lemma 2, the sequence of the ellipsoids does not shrink to a point a.s. $\square$

2. THE OBE-ABE ALGORITHM

In this section, we introduce the OBE-ABE algorithm whose estimator is a.s. consistent and/or p. consistent under various conditions. The error sequence $\{v_n\}$ is assumed stationary for a.s. consistency. Stationarity and the independence between $\{u_n\}$ and $\{v_n\}$ are not necessary for p. consistency [6].

Theorem 2 Assume that $\{v_n\}$ is independent of $\{u_n\}$. If PE holds, UCT holds, and $\{v_n\}$ is mixing, then the estimator of the OBE-ABE algorithm is a.s. consistent.

In the above theorem, mixing of $\{v_n\}$ can be relaxed to asymptotic independence of $\{v_n\}$ for p. consistency. Further, if $\{v_n\}$ or $\{u_n\}$ or both are continuously-distributed random sequences, then, the mixing condition can be relaxed to an ergodic condition for a.s. consistency. The following theorem validates this assertion.

Theorem 3 Assume that $\{v_n\}$ is independent of $\{u_n\}$. If $\{v_n\}$ or $\{u_n\}$ or both are continuously distributed random sequences, PE holds, UCT holds, and $\{v_n\}$ is ergodic, then the estimator of the OBE-ABE algorithm is a.s. consistent.

For p. consistency, ergodicity of $\{v_n\}$ and the independence of $\{u_n\}$ and $\{v_n\}$ are not required in Theorem 3. The following theorem is an important special case of Theorem 2.

Corollary 1 Assume that $\{v_n\}$ is independent of $\{u_n\}$. If PE holds, UT holds, and $\{v_n\}$ is independently identically distributed (i.i.d.), then the estimator of the OBE-ABE algorithm is a.s. consistent.

Proof: Since UT and UCT imply each other with the i.i.d. assumption, and i.i.d. implies mixing, a.s. consistency follows immediately from Theorem 2. $\square$

The following corollaries assert the a.s. consistency of the OBE-ABE algorithm when certain distributions of v_n are known.

Corollary 2 Assume that $\{v_n\}$ is stationary and independent of $\{u_n\}$. If $\{v_n\}$ is uniformly distributed and PE holds, then the estimator of the OBE-ABE algorithm is a.s. consistent.

Corollary 3 Assume that $\{v_n\}$ is stationary and independent of $\{u_n\}$. If $\{v_n\}$ is binary Bernoulli distributed with $P(v_n = \sqrt{\gamma_*}) > \delta$ or $P(v_n = -\sqrt{\gamma_*}) > \delta$, $\forall n$, and PE holds, then the estimator of the OBE-ABE algorithm is a.s. consistent.

Note that all the theorems in this section are also valid for conventional OBE algorithms (and the EPA algorithm) with exact bounds since, for the OBE-ABE algorithm, $\gamma_n \rightarrow \gamma_*$ as $n \rightarrow \infty$ [6]. Please refer to [6, 9] for details.

3. SUB-OBE-ABE ALGORITHM

In this section, we introduce a modified OBE-ABE algorithm (the Sub-OBE-ABE algorithm) in which an $\mathcal{O}(m)$ checking for innovation is employed.

A modification of conventional OBE algorithms to achieve $\mathcal{O}(m)$ checking can be found in [2, 3]. The modified OBE algorithms (Sub-OBE algorithm) in [2, 3] are actually $\mathcal{O}(\rho m^2)$ complexity with ρ indicating the fraction of data found to be innovative. An $\mathcal{O}(m)$ check for innovation is used which is similar to the check in [1]. Hence, the Sub-OBE algorithm in [2, 3] is an $\mathcal{O}(m)$ algorithm for small m since ρ is typically near 0.1 for a uniformly distributed $\{v_n\}$.

However, the Sub-OBE checking cannot be applied directly to the OBE-ABE algorithm, since the blind error-bound estimation procedure will fail [6]. Hence, the Sub-OBE-ABE algorithm is incorporated with an $\mathcal{O}(m)$ checking formula for innovation: $c'_n = m\gamma_n - m\epsilon_n^2 - \kappa_{n-1} \mathbf{x}_n^T \mathbf{x}_n / d_{n-1} < 0$, where d_n has the recursion: $d_n = (1 - \lambda_n)d_{n-1} + \lambda_n \mathbf{x}_n^T \mathbf{x}_n$. λ_n in this recursion is the data weight. The Sub-OBE-ABE algorithm is compared to the OBE-ABE algorithm with regard to convergence speed and computational complexity in Section 5.

All the theorems in the previous section are valid for the Sub-OBE-ABE algorithm under the same conditions [6].

4. ADAPTIVE SUB-OBE-ABE

Adaptive OBE algorithms have been shown to have superior tracking capability to RLS, LMS, and their variants [2, 11]. In [11], Rao and Huang modified the OBE algorithm by resetting κ_{n-1} , whenever $\kappa_n < 0$, to a value $\zeta + K_1$ or $\zeta + K_2$, where ζ is set to 1, and K_1 , and K_2 are two positive values depending on the algorithm's parameters and data at time n . This is equivalent to the resetting of κ_{n-1} to a value greater than 1 whenever $\kappa_n < 0$.

In [2], Deller *et al.* proposed numerous methods for modifying OBE algorithms to be adaptive to time-varying environments. Among those methods, best tracking performance results from the *selective forgetting* method which back-rotates previously accepted data sets when $\kappa_n < 0$, until $\kappa_n > 0$.

In this paper, we propose an *adaptive* Sub-OBE-ABE algorithm by modifying Sub-OBE-ABE to include resetting of $\kappa_n = 0.1$ whenever $\kappa_n < 0$. Simulations show that this algorithm has excellent tracking performance in slow and fast time-varying environments.

5. SIMULATIONS

In this section, we describe simulations to support the theorems and adaptive Sub-OBE-ABE algorithm proposed in the previous sections.

The simulations for the consistent convergence of the OBE-ABE algorithm in a non-zero-mean i.i.d. error environment are found in [9]. Here, we compare the simulation results of the OBE-ABE algorithm with those of conventional OBE (SM-SA) algorithm using an AR(12) model with correlated non-zero-mean error sequence $\{v_n\}$: $y_n = a_{1*}y_{n-1} + a_{2*}y_{n-2} + \dots + a_{12*}y_{n-12} + v_n$; where $a_{1*} = -0.1$, $a_{2*} = 0.9175$, $a_{3*} = -0.191$, $a_{4*} = -0.2253$, $a_{5*} = 0.2601$, $a_{6*} = 0.0046$, $a_{7*} = -0.0367$, $a_{8*} = -0.0209$, $a_{9*} = -0.0082$, $a_{10*} = 0.0095$, $a_{11*} = -0.0052$, and $a_{12*} = -0.0041$, and v_n is a correlated non-zero-mean sequence generated by a correlated sequence $\{w_n\}$: $v_n = \{1, \text{ if } w_n > -1; -1, \text{ otherwise}\}$. w_n is generated by: $w_n = -0.8w_{n-1} + z_n$, where $z_n \sim U(-1, 1)$ is i.i.d. Both algorithms run with an overestimated bound $\gamma_0 = 1.5$ since the true error bound of $\{v_n\}$ ($\gamma_* = 1$) is assumed unknown. As seen in Figs. 1 and 2, the OBE-ABE algorithm, within 2000 steps, converges with 7.5% of data being selected while the OBE algorithm, with 8% of data being selected, does not converge well to the true parameter $a_{1*} = -0.1$.

Simulations also show that, for both algorithms to achieve same desired volume of ellipsoid, the Sub-OBE-ABE algorithm requires fewer steps and fewer selected data than those of Sub-OBE algorithms or OBE algorithms.

Simulations show that the convergence speed of the Sub-OBE-ABE algorithm and the OBE-ABE algorithm are comparable while the Sub-OBE-ABE algorithm selects fewer data. Figures 3 and 4 show the simulation results in the same correlated-error noise environment except the model is an AR(3) with $a_{1*} = 2$, $a_{2*} = -1.48$, $a_{3*} = 0.34$. In this case, the Sub-OBE-ABE algorithm uses, on average, 6% of the data to update the estimator while the OBE-ABE algorithm uses 16%. Other simulations, in the case of i.i.d. and uniformly distributed $\{v_n\}$, show that the Sub-OBE-ABE algorithm uses, on average, 3.5% of the data to update the estimator while the OBE-ABE algorithm uses 6%. Simulations also show that the convergence performance and computational efficiency of Sub-OBE-ABE are not affected by the order (m) of system.

Both Sub-OBE-ABE and OBE-ABE are more robust to a noisy environment than conventional OBE algorithms. Please refer to [6, 9] for more details.

To show the tracking performance of the Sub-OBE-ABE algorithm, let us consider a time-varying AR(2) model $y_n = a_{1*}y_{n-1} - 0.68y_{n-2} + v_n$ in which a_{1*} varies between -1.6 and 1.6 (as in [2]). This is equivalent to varying the system's conjugate poles $0.8 \pm j0.2$ to and from $-0.8 \pm j0.2$. The variations of a_{1*} (abruptly or gradually) are shown as dashed lines in Figs. 5 and 6. The error sequence $\{v_n\}$ is i.i.d. and uniformly distributed on $[-1, 1]$.

Although "non-adaptive" OBE algorithms often have good tracking capability, they eventually lose tracking in time-varying environments as shown in [2, 11]. However, in the same environments or faster-changing environments, the adaptive Sub-OBE-ABE algorithm ($\gamma_0 = 1.5$, $\epsilon = 0.02$, $M = 70$) keeps track of the variations very well with no more than 7.5% of data selected (ρ in each figure) as shown in Figs. 5 and 6 (not shown here for slow system). Comparing these diagrams with those in [2, 11] shows that the adaptive Sub-OBE-ABE algorithm has comparable or better tracking capability.

6. CONCLUSION

This paper presents an OBE-ABE algorithm, Sub-OBE-ABE algorithm, and an adaptive version of the two algorithms. Included are sufficient conditions for a.s. and p. consistent convergence of all OBE-type algorithms (and the EPA algorithm [12]) under various conditions (i.i.d., correlated, or non-stationary) for the ARX model. A necessary condition for the shrinking of ellipsoids of OBE algorithms (hence, the OBE-ABE and Sub-OBE-ABE algorithms) to a point a.s. is also included. Simulations demonstrate that, if the noise bound γ_* of random sequence $\{v_n\}$ is unknown *a priori*, the new algorithms proposed in this paper are superior to any conventional OBE algorithm with respect to convergence, speed of convergence, computational efficiency, robustness to measurement noise, and tracking capability in slow and fast time-varying environments.

REFERENCES

- [1] Dasgupta, S., and Y. F. Huang, "Asymptotically convergent modified recursive least squares with data dependent updating and forgetting factor for systems with bounded noise," *IEEE Trans. Info. Theory*, vol. 33, pp. 383-392, 1987.
- [2] Deller Jr., J. R., M. Nayeri and S. F. Odeh, "Least-square identification with error bounds for real-time signal processing and control," *Proc. IEEE*, vol. 81, pp. 815-849, 1993.
- [3] Deller Jr., J. R. and S. F. Odeh, "Adaptive set-membership identification in $\mathcal{O}(m)$ time for linear-in-parameters models," *IEEE Trans. on Sig. Proc.*, vol. 41, No. 5, May, 1993.
- [4] Deller Jr., J. R., M. Nayeri, and M. S. Liu, "Unifying the landmark developments in optimal bounding ellipsoid identification," *Int. J. of Adapt. Contr. and Sig. Proc.*, vol. 8, pp. 43-60, 1994.
- [5] Fogel, E., and Y. F. Huang, "On the value of information in system identification - bounded noise case," *Automatica*, vol. 18, pp. 229-238, 1982.
- [6] Lin, T. M., *OBE Algorithms with Automatic Bound Estimation* (Ph.D. dissertation), Michigan State Univ., Dec., 1996.
- [7] Liu, M. S., *Development and Analysis of an Optimal Bounding Ellipsoid Algorithm Using Stochastic Approximation* (Ph.D. dissertation), Michigan State University, 1993.
- [8] Liu, M.S., M. Nayeri, and J.R. Deller Jr., "Do interpretable optimal bounding ellipsoid algorithms converge? Part II - OBE vs. RLS: Clearing the smoke," *Proc. 10th IFAC Symp. on System Identification (SYSID '94)*, July, vol. 3, pp. 395-400, Copenhagen, Denmark, 1994.
- [9] Nayeri, M., T. M. Lin, and J. R. Deller Jr., "Blind error-bound estimation for OBE identification," *Proc. 34th Ann. Allerton Conf. on Commun., Contr. and Comput.*, Univ. of Illinois, Urbana-Champaign, Oct 2-4, 1996 (in press).
- [10] Nayeri, M., M. S. Liu, and J. R. Deller Jr., "Do interpretable optimal bounding ellipsoid algorithms converge? Part I - The long-awaited set-convergence proof," *Proc. 10th IFAC Symp. on System Identification (SYSID '94)*, July, vol. 3, pp. 389-394, Copenhagen, Denmark, 1994.
- [11] Rao, A. K., and Y. F. Huang, "Tracking characteristics of an OBE parameter estimation algorithm," *IEEE Trans. Sig. Proc.*, vol. 41, pp. 1140-1148, 1993.
- [12] Veres, S.M., and J.P. Norton, "Structure selection for bounded-parameter models: Consistency conditions and selection criterion," *IEEE Trans. on Auto. Contr.*, vol. 36, pp. 474-481, 1991.

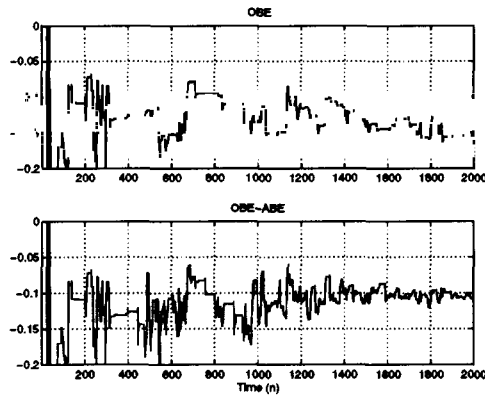


Figure 1. Estimates a_1 of OBE (SM-SA) ($\gamma_0 = 1.5$) and Sub-OBE-ABE ($\gamma_0 = 1.5$, $\epsilon = 0$, $M = 50$) in a correlated-error environment where $a_{1*} = -0.1$.

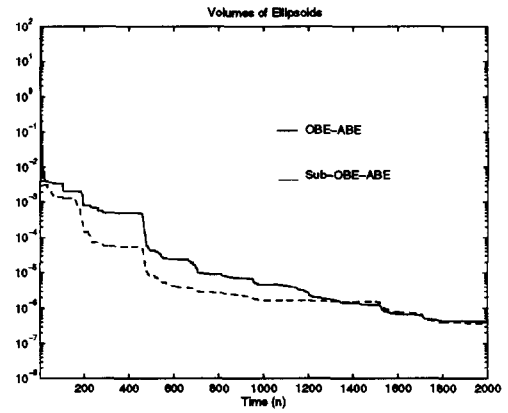


Figure 4. Volumes of ellipsoids of OBE-ABE and Sub-OBE-ABE ($\gamma_0 = 1.5$, $\epsilon = 0$, $M = 50$) in a correlated-error environment.

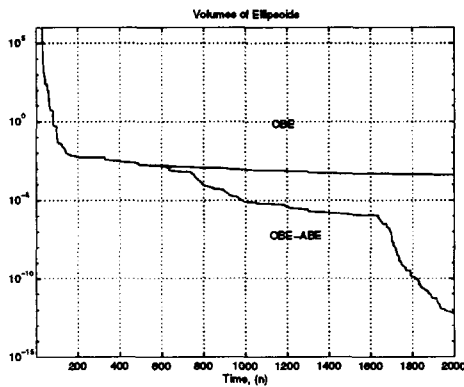


Figure 2. Volumes of ellipsoids of OBE (SM-SA) ($\gamma_0 = 1.5$) and Sub-OBE-ABE ($\gamma_0 = 1.5$, $\epsilon = 0$, $M = 50$) in a correlated-error environment.

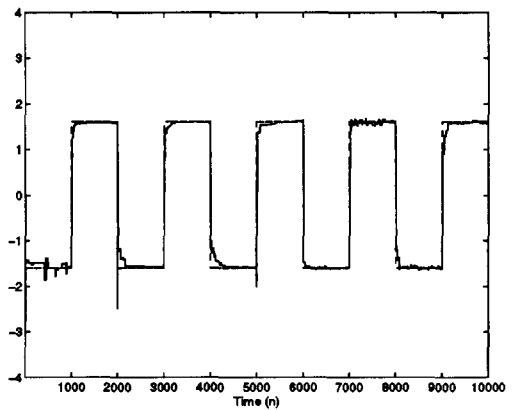


Figure 5. Result of using adaptive Sub-OBE-ABE on an abruptly varying system. $\rho = 3.5\%$.

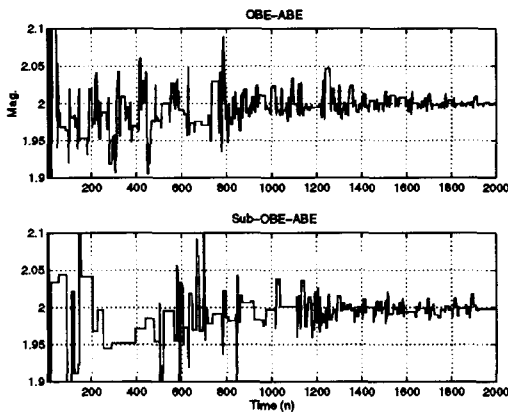


Figure 3. Estimates a_1 of OBE-ABE and Sub-OBE-ABE ($\gamma_0 = 1.5$, $\epsilon = 0$, $M = 50$) in a correlated-error environment where $a_{1*} = 2$.

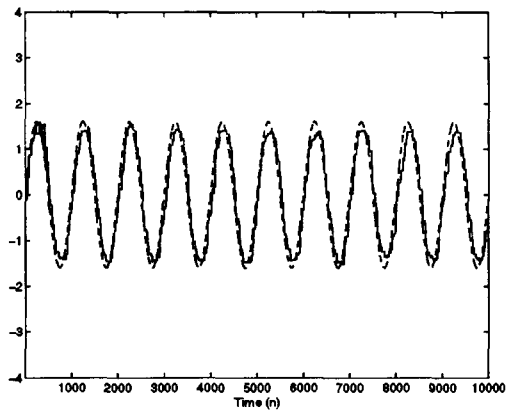


Figure 6. Result of using adaptive Sub-OBE-ABE on a gradually varying system. $\rho = 7.5\%$.