ON A PERTURBATION APPROACH FOR THE ANALYSIS OF STOCHASTIC TRACKING ALGORITHMS

E. Moulines

ENST, 46, rue Barrault, 75634 Paris

P. Priouret, R. Aguech

Labo, de Probabilite, Univ. Paris VI, 75252 Paris

ABSTRACT

In this paper, a perturbation expansion technique is introduced to decompose the tracking error of a general adaptive tracking algorithm in a linear regression model. This method allow to obtain tracking error bound but also tight approximate expressions for the moments of the tracking error. These expressions allow to evaluate, both qualitatively and quantitatively, the impact of several factors on the tracking error performance which have been overlooked in previous contributions.

1. INTRODUCTION

An important issue in system identification, signal processing, automatic control is that of tracking the parameter variations in a linear regression model

$$y_t = \phi_t^T \theta_t + v_t \, ; \;\; t \geq 0 \tag{1}$$

where $\{y_t\}_{t\geq 0}$ and $\{v_t\}_{t\geq 0}$ are respectively the scalar observation and noise, $\{\phi_t\}_{t\geq 0}$ and $\{\theta_t\}_{t\geq 0}$ are the d-dimensional stochastic regressor and the unknown time-varying parameter. This model encompasses many different applications, including channel equalization, time delay estimation and echo cancelation [5]. In the sequel it is assumed that the parameter variation obeys

$$\theta_{t+1} = \theta_t + w_{t+1} \tag{2}$$

where w_{t+1} is referred to the lag-noise. To track the variations of the parameter, it is customary to use a recursive algorithm for updating an estimate $\hat{\theta}_t$ of the parameter (see [6, 5] and the references therein). Most of these algorithms can be put in the form

$$\hat{\theta}_{t+1} = \hat{\theta}_t + \mu L_t (y_t - \phi_t^T \hat{\theta}_t). \tag{3}$$

where μ is referred to as the adaptation step-size and L_t is a random vector, which can be chosen in a number of different ways. There is a vast literature on the analysis of algorithms of type (3). In most contributions, the main goal is to obtain bounds on the tracking errors. Results in that directions have been obtained in [6, 4, 3]. In this contribution a different approach is pursued. Our goal is to obtain explicit expression and not only bounds for the tracking error. To that purpose, we will use a technique, referred to as 'perturbation expansion', consisting in getting approximations (3) by nested processes, with much simpler structure than the original error process. This particular decomposition enables the computation of explicit expressions for the moments and other related quantities.

2. PERTURBATION EXPANSION: OUTLINES OF THE METHOD

From (1) and (3), we can write

$$\tilde{\theta}_{t+1} = (I - \mu L_t \phi_t^T)\tilde{\theta}_t + \mu L_t v_t - w_{t+1}, \tag{4}$$

where $\tilde{\theta}_t = \hat{\theta}_t - \theta_t$ is the weight-error vector. Since this equation is linear, $\tilde{\theta}_{t+1}$ can be decomposed as

$$\begin{aligned}
\tilde{\theta}_t &= {}^u\tilde{\theta}_t + \mu\, {}^v\tilde{\theta}_t + {}^w\tilde{\theta}_t, \\
{}^u\tilde{\theta}_{t+1} &= (I - \mu L_t \phi_t^T)\, {}^u\tilde{\theta}_t \quad {}^u\tilde{\theta}_0 = \tilde{\theta}_0 = -\theta_0, \\
{}^v\tilde{\theta}_{t+1} &= (I - \mu L_t \phi_t^T)\, {}^v\tilde{\theta}_t + L_t v_t, \;\; {}^v\tilde{\theta}_0 = 0 \\
{}^w\tilde{\theta}_{t+1} &= (I - \mu L_t \phi_t^T)\, {}^w\tilde{\theta}_t - w_{t+1}, \;\; {}^w\tilde{\theta}_0 = 0
\end{aligned}$$

$\{{}^u\tilde{\theta}_t\}$ is a transient term, reflecting the way the successive estimates of the regression coefficients forget the initial conditions. $\{{}^v\tilde{\theta}_t\}$ accounts for the errors introduced by the measurement noise, $\{v_t\}$; similarly, $\{{}^w\tilde{\theta}_{t+1}\}$ accounts for the errors associated lag-noise $\{w_t\}$. According to these definitions, ${}^v\tilde{\theta}_t$ and ${}^w\tilde{\theta}_t$ obey an inhomogeneous stochastic recurrence equation

$$\begin{aligned}
\delta_{t+1} &= (I - \mu F_t)\delta_t + \xi_t, \quad \delta_0 = 0 &(5)\\
&= \sum_{s=0}^{t} \Phi(t,s)\xi_s &(6)
\end{aligned}$$

where $\{F_t\}_{t\geq 0}$ matrix valued random process, $\{\xi_t\}_{t\geq 0}$ is a $(d \times 1)$ vector-valued random process, and $\Phi(t,s)$ is defined as

$$\Phi(t,s) = \begin{cases} (I - \mu F_t)(I - \mu F_{t-1}) \cdots (I - \mu F_{s+1}) & t > s \\ I & t = s \\ 0 & \text{otherwise} \end{cases}$$

Here, the dependence of δ_t upon the step-size μ is implicit. Eqs (2) and (2) may be rewritten as (5) with $F_t = L_t \phi_t^T$ and

$$\xi_t = L_t v_t \text{ measurement noise, } \xi_t = -w_{t+1} \text{ lag noise} \tag{7}$$

Applied to the recurrence equation (5), the whole procedure goes as follows. Denote $\bar{F}_t = E(F_t)$ and $Z_t = \bar{F}_t - F_t$. We may decompose $(I - \mu F_t)$ according to

$$I - \mu F_t = (I - \mu \bar{F}_t) + \mu Z_t. \tag{8}$$

Now, decompose the recurrence equations (5) into two separate recursions:

$$\begin{aligned}
J_{t+1}^{(0)} &= (I - \mu \bar{F}_t) J_t^{(0)} + \xi_t, \quad J_0^{(0)} = 0 &(9)\\
H_{t+1}^{(0)} &= (I - \mu F_t) H_t^{(0)} + \mu Z_t J_t^{(0)}, \quad H_0^{(0)} = 0 &(10)\\
\delta_t &= J_t^{(0)} + H_t^{(0)}. &(11)
\end{aligned}$$

According to (9), $J_t^{(0)}$ satisfy a *deterministic* inhomogeneous first-order linear difference equation:

$$J_{t+1}^{(0)} = \sum_{s=0}^{t} \psi(t,s)\xi_s \tag{12}$$

where, as above,

$$\psi(t,s) = \begin{cases} (I-\mu\bar{F}_t)(I-\mu\bar{F}_{t-1})\cdots(I-\mu\bar{F}_{s+1}) & t > s \\ I & t = s \\ 0 & \text{otherwise} \end{cases}$$

Under appropriate assumptions on the matrix valued sequences $\{F_t\}_{t\geq 0}$ and on the excitation $\{\xi_t\}$, it will be shown that, for some $p > 0$ there exists a constant $C < \infty$ and $\mu_0 > 0$ such that for all $0 < \mu \leq \mu_0$

$$\sup_{t\geq 0} \|J_t^{(0)}\|_p \leq C/\sqrt{\mu} \text{ and } \sup_{t\geq 0} \|H_t^{(0)}\|_p \leq C \tag{13}$$

where $C < \infty$ is a constant depending on $\{F_t\}$ and $\{\xi_t\}$ (see below). Thus, $J_t^{(0)}$ may be considered as the leading term in the expansion, while $H_t^{(0)}$ may be seen as a correction term. The same procedure can be iterated to obtain approximations of increased accuracy. For that purpose, it suffices to decompose (10) using (8), and iterate that decomposition up to order $n > 1$. Using this technique, δ_t may be written as

$$\delta_t = J_t^{(0)} + J_t^{(1)} + \cdots + J_t^{(n)} + H_t^{(n)}, \tag{14}$$

where the processes $J_t^{(r)}$, $0 \leq r \leq n$ and $H_t^{(n)}$ are respectively defined as

$$J_{t+1}^{(0)} = (I-\mu\bar{F}_t)J_t^{(0)} + \xi_t \; ; \;\; J_0^{(0)} = 0$$
$$J_{t+1}^{(r)} = (I-\mu\bar{F}_t)J_t^{(r)} + \mu Z_t J_t^{(r-1)} \; ; \;\; J_t^{(r)} = 0, \; 0 \leq t < r$$
$$H_{t+1}^{(n)} = (I-\mu F_t)H_t^{(n)} + \mu Z_t J_t^{(n)} \; ; \;\; H_t^{(n)} = 0, \; 0 \leq t < n$$

The processes $J_t^{(r)}$ depend linearly on ξ_t and polynomially in the error $Z_t = \bar{F}_t - F_t$. It is thus feasible (examples are given below) to compute the joint moments of these processes, and to obtain expressions for the moments of $\tilde{\delta}_t^{(n)} = J_t^{(0)} + \cdots + J_t^{(n)}$. The residual term $H_t^{(n)}$ is, under appropriate conditions, uniformly bounded, *i.e.*, there exists some constant $C < \infty$ and $\mu_0 > 0$, such that, for all $0 < \mu \leq \mu_0$, we have

$$\sup_{t\geq 0} \|H_t^{(n)}\|_p \leq C\mu^{n/2}. \tag{15}$$

Upper bounds for the constant C are given below.

3. MAIN RESULTS

3.1. Notations and Definitions

To develop a useful theory, one typically needs to impose assumptions on $\{F_t\}$ and $\{\xi_t\}$. It is not difficult to guess that typical assumptions are of the kind **(i)** *restriction on the tail of the distribution*, e.g. existence of moments and / or exponential moments and **(ii)** *restriction on the dependence among the* $\{F_t\}$ *and the* $\{\xi_t\}$. The latter conditions are referred to as *mixing conditions*. To make more precise statements, some notations and definitions are in order.

Definition Let $q \geq 1$ and let $X = \{X_n\}_{n\geq 0}$ be a $(l \times 1)$ vector-valued process. Let $\delta = (\delta(r))_{r\in\mathbb{N}}$ a sequence decreasing to zero at infinity. $X = \{X_n\}_{n\geq 0}$ is called (δ, q)-weak dependent, if there exists finite constants $C = \{C_1, \cdots, C_q\}$, such that for any $1 \leq m < s \leq q$, any m-tuple $t_1, \cdots, t_m$ and any $s-m$-uple $t_{m+1}, \cdots, t_s$, with $t_1 \leq \cdots \leq t_m < t_m + r \leq t_{m+1} \leq \cdots t_s$, it holds

$$\sup_{i_1,\cdots,i_s} \left|\text{cov}(X_{t_1,i_1}\cdots X_{t_m,i_m}, X_{t_{m+1},i_{m+1}}\cdots X_{t_s,i_s})\right| \leq C_s\delta(r)$$

where $X_{n,i}$ denotes the i-th component of X_n. As shown in [8], weak-mixing processes encompass a large class of models and in particular strongly mixing processes. The following result plays a key role in the sequel

Proposition 1 *Let* $p \geq 1$ *and* $n \in \mathbb{N}$. *Let* $G = \{G_t\}_{t\geq 0}$ *be a* $(d \times d)$ *zero-mean matrix-valued process. Assume that* $(\delta, p(n+2))$ *weak-dependent and that*

$$\sum (r+1)^{p(n+2)/2-1}\delta(r) < \infty. \tag{16}$$

Then, there exists a finite constant $D_{p,n}(G)$*, such that for all* $j \in \{1, \cdots, n\}$ *and all* $0 \leq s \leq t < \infty$*, we have*

$$\left\| \sum_{s\leq i_1<\cdots<i_j\leq t} G_{i_1}\cdots G_{i_j} \right\|_{pn/j} \leq D_{p,n}(G)(t-s)^{j/2}. \tag{17}$$

Let $\mathcal{B}$ be a subfield of the basic probability space $(\Omega, \mathcal{A})$.

- Let $q \geq p > 0$ be two real numbers. For any $\eta = \{\eta_k\}_{k\geq 0}$ $d \times 1$ vector-valued sequence, define $\mathcal{N}(p, q; \mathcal{B})$ as the set

$$\Big\{\eta : \left\|\sum_{k=s}^{t} G_k\eta_k\right\|_p \leq \rho_{p,q}(\eta)\left(\sum_{k=s}^{t}\|G_k\|_q^2\right)^{1/2}$$
$$\forall\; 0 \leq s \leq t \; \forall\; G = \{G_k\}_{k\geq 0}\; (1\times d) \;\in L_q(\Omega, \mathcal{B}, P)\Big\}$$

Conditions upon which a process $\{\xi_k\}$ belongs to the set $\mathcal{N}(p, q; \mathcal{B})$ and expressions of the constants $\rho_{p,q}(\xi)$ are given in the extended version of this paper.

3.2. The Main Results

We may now formulate the central results of our contribution. The first result gives condition upon which $J_s^{(r)}$ is uniformly bounded in L_p and provides an expression for that bound.

Theorem 1 *Assume that, for some integer* n *and real numbers* $q \geq p \geq 2$*:*

- **(i)** $F = \{F_t\}_{t\geq 0}$ *is averaged exponentially stable,*
- **(ii)** F *is* $(\delta, q(n+2))$ *weakly dependent, and*

$$\sum (r+1)^{(q(n+2)/2-1}\delta(r) < \infty.$$

- **(iii)** $\xi = \{\xi_t\} \in \mathcal{N}(p, q; F)$

Then, there exists a constant $K < \infty$ *(depending on* F *and on the numerical constants* p, q, n, μ_0, β but not *on* $\{\xi_t\}$ *or on the stepsize parameter* μ*), such that for all* $0 < \mu \leq \mu_0$*, for all* $0 \leq r \leq n$

$$\sup_{s\geq 1} \|J_s^{(r)}\|_p \leq K\; \rho_{p,q}(\xi)\mu^{(r-1)/2}. \tag{18}$$

where $\rho_{p,q}(\xi)$ *is the constant defined in ().*

Here $\mathcal{N}(p,q;F)$ is a shorthand notation for $\mathcal{N}(p,q;\mathcal{M}_0^\infty(F))$. To complete our program, we need to bound the remainder term $H_s^{(n)}$. As shown below, under appropriate conditions, $H_s^{(n)}$ is uniformly bounded in L_q as soon as $J_s^{(n+1)}$ is L_r stable (for sufficiently large r) and the bound for $\sup_s \|H_s^{(n)}\|_q$ is proportional to $\sup_s \|J_s^{(n+1)}\|_r$.

Theorem 2 *Let $p \geq 2$ and let $a,b,c > 0$ such that $1/a + 1/b + 1/c = 1/p$. Let $n \in \mathbb{N}$. Assume that*

- *$\{F_t\}$ is L_a-exponentially stable,*
- *$\sup_{s\geq 0} \|Z_s\|_b < \infty$ and,*
- *$\sup_{s\geq 0} \|J_s^{(n+1)}\|_c < \infty$.*

Then, there exists a constant $K' < \infty$ (depending on the numerical constants a,b,c,β,μ_0,n and on the process $\{F_t\}$ but not on the process $\{\xi_t\}$ or on the stepsize parameter μ), such that for all $0 < \mu \leq \mu_0$,

$$\sup_{s\geq 0} \|H_s^{(n)}\|_p \leq K' \sup_{s\geq 0} \|J_s^{(n+1)}\|_c. \tag{19}$$

4. PERFORMANCE OF ADAPTIVE TRACKING ALGORITHMS

A number of useful error bounds or expressions can be drawn from the results derived in the previous section. We use the notations introduced in section 1. Let n be a positive integer and p be a real number $p \geq 2$. Finally, let a,b,c,d be positive numbers such that $a^{-1}+b^{-1}+c^{-1} = p^{-1}$ and $d \geq c$. Denote: $F_t = L_t \phi_t^T$.

- **(H1)** $F = \{F_t\}$ is L_a and averaged exponentially stable,
- **(H2)** F is $(\delta, d(n+2))$ weak dependent; in addition,

$$\sup_{t\geq 0} \|F_t - E(F_t)\|_b < \infty,$$
$$\sum_{r=0}^{\infty} (r+1)^{d(n+2)/2-1} \delta(r) < \infty.$$

- **(H3)** $\{L_t v_t\} \in \mathcal{N}(c,d;F)$ and $\{w_t\} \in \mathcal{N}(c,d;F)$.

The tracking error may be expanded as $\tilde{\theta}_t =^u \tilde{\theta}_t + \mu^v \tilde{\theta}_t +^w \tilde{\theta}_t$, where ${}^u\tilde{\theta}_t$, ${}^v\tilde{\theta}_t$ and ${}^w\tilde{\theta}_t$ are respectively defined in (2), (2) and (2). The terms ${}^v\tilde{\theta}_t$ and ${}^w\tilde{\theta}_t$ may further be decomposed as

$${}^v\tilde{\theta}_t = \sum_{k=0}^{r_v} {}^vJ_t^{(k)} + {}^vH_t^{(r_v)}$$
$${}^w\tilde{\theta}_t = \sum_{k=0}^{r_w} {}^wJ_t^{(k)} + {}^wH_t^{(r_w)}$$

where r_v and r_w are two integers (not necessarily equal) such that $0 \leq r_v \leq n-1$ and $0 \leq r_w \leq n-1$. Theorems 1, 2 show that

Proposition 2 *Assume that (H1-H2-H3) hold, and let $r_v, r_w \in \{0,\cdots,n-1\}$. Then, there exists a constant $K < \infty$ (depending upon a,b,c,d,n,β,μ_0 but not on $\{L_t v_t\}$ and $\{w_t\}$), such that, for all $\mu \in (0,\mu_0]$ and all $t \geq 0$*

$$\|\tilde{\theta}_t - \mu \sum_{k=0}^{r_v} {}^vJ_t^{(k)} - \sum_{k=0}^{r_w} {}^wJ_t^{(k)}\|_p \leq$$
$$K\left(\rho_{c,d}(Lv)\mu^{r_v/2+1} + \rho_{c,d}(w)\mu^{r_w/2}\right) + \|\Phi(t,0)\tilde{\theta}_0\|_p.$$

5. A WORKED-OUT EXAMPLE

In this section, approximate expressions of the tracking error covariance matrix for the LMS algorithm are derived. To illustrate our findings, it is shown in this section that first order approximation of the tracking error covariance may fail, in certain situation, to capture the behavior of the algorithm.

- **(M1)** $\{\phi_t\}_{t\geq 0}$ is VAR process
$$\phi_{t+1} = \kappa \phi_t + u_{t+1} \tag{20}$$
where κ $(-1 < \kappa < 1)$ is a scalar, $\{u_t\}_{t\in\mathbb{Z}}$ is i.i.d Gaussian with zero-mean and covariance matrix $\sigma_u^2 I$.
- **(M2)** $\{v_t\}_{t\geq 0}$ and $\{w_t\}_{t\geq 0}$ are i.i.d, with bounded moments of order r, where $r > 2$. Moreover: $E(w_0 w_0^T) = \gamma^2 I$.
- **(M3)** $\mathcal{M}_0^\infty(v)$, $\mathcal{M}_0^\infty(\phi)$ and $\mathcal{M}(\theta)$ are independent.

It may be shown that (M1–M3) implies (H1–H3). Because $\{w_t\}_{t\geq 0}$ and $\{v_t\}_{t\geq 0}$ are independent, the processes $\{{}^v\tilde{\theta}_t\}_{t\geq 0}$ and $\{{}^w\tilde{\theta}_t\}$ are uncorrelated. Thus,

$$\Gamma = \lim_{t\to\infty} E(\tilde{\theta}_t \tilde{\theta}_t^T) = {}^v\Gamma\mu^2 + {}^w\Gamma$$

where ${}^v\Gamma = \lim_{t\to\infty} E({}^v\tilde{\theta}_t^v\tilde{\theta}_t^T)$ and ${}^w\Gamma = \lim_{t\to\infty} E({}^w\tilde{\theta}_t^w\tilde{\theta}_t^T)$. We wish to obtain approximate expressions for ${}^v\Gamma$ and ${}^w\Gamma$, denoted ${}^v\bar{\Gamma}$ and ${}^w\bar{\Gamma}$ such that, for all $\mu \in (0,\mu_0]$

$$|{}^v\Gamma -^v \bar{\Gamma}| \leq K\mu^{1/2} \text{ and } |{}^w\Gamma -^w \bar{\Gamma}| \leq K\gamma^2\mu^{1/2}$$

where $K < \infty$ is some constant. To that purpose, we expand ${}^v\tilde{\theta}_t$ and ${}^w\tilde{\theta}_t$ to the second-order:

$${}^v\tilde{\theta}_t = {}^vJ_t^{(0)} + {}^vJ_t^{(1)} + {}^vJ_t^{(2)} + {}^vH_t^{(2)},$$
$${}^w\tilde{\theta}_t = {}^wJ_t^{(0)} + {}^wJ_t^{(1)} + {}^wJ_t^{(2)} + {}^wH_t^{(2)}$$

Under the stated assumptions, it follows from theorems 1 and 2 that, there exists some constant $C < \infty$, such that for all $\mu \in (0,\mu_0]$, we have

$$\sup_{t\geq 0} \left| E({}^vJ_t^{(1)}({}^vJ_t^{(2)} + {}^vH_t^{(2)})^T \right| \leq C\mu^{1/2}$$
$$\sup_{t\geq 0} \left| E({}^wJ_t^{(1)}({}^wJ_t^{(2)} +^w H_t^{(2)})^T \right| \leq C\gamma^2\mu^{1/2}$$

It remains to evaluate $\lim_{t\to\infty} E({}^vJ_t^{(0)\,v}J_t^{(i)})$, $\lim_{t\to\infty} E({}^wJ_t^{(0)\,w}J_t^{(i)})$, $i = 0,1,2$ and $E({}^vJ_t^{(1)\,v}J_t^{(1)})$ and $E({}^wJ_t^{(1)\,w}J_t^{(1)})$. Denote: $\alpha = \sigma_u^2/(1-\kappa^2)$. Tedious but straightforward calculations show that

$$\lim_{t\to\infty} E({}^vJ_t^{(0)\,v}J_t^{(0)\,T}) = \frac{\sigma_v^2}{\mu(2-\mu\alpha)},$$
$$\lim_{t\to\infty} E({}^vJ_t^{(0)\,v}J_t^{(1)\,T}) = -\frac{\kappa^2\sigma_v^2(d+1)\alpha}{2(1-\kappa^2)} I + O(\mu),$$
$$\lim_{t\to\infty} E({}^vJ_t^{(0)\,v}J_t^{(2)\,T}) = \frac{\kappa^2\sigma_v^2\alpha(d+1)\alpha}{4(1-\kappa^2)} I + O(\mu),$$
$$\lim_{t\to\infty} E({}^vJ_t^{(1)\,v}J_t^{(1)\,T}) = \frac{(1+\kappa^2)\sigma_v^2\alpha(d+1)}{4(1-\kappa^2)} + O(\mu)$$

yielding the following expression for ${}^v\bar{\Gamma}$

$${}^v\bar{\Gamma} = \frac{\sigma_v^2}{2\mu} I + \alpha\sigma_v^2 \frac{d+2}{4} I + O(\mu) \tag{21}$$

It is worthwhile to note that the first order correction does not depend upon the autoregressive coefficient κ. Similarly, it can be shown that

$$\lim_{t\to\infty} E({}^wJ_t^{(0)}\,{}^wJ_t^{(0)\,T}) = \frac{\gamma^2}{\mu\alpha(2-\mu\alpha)} I,$$
$$\lim_{t\to\infty} E({}^wJ_t^{(0)}\,{}^wJ_t^{(1)\,T}) = 0,$$
$$\lim_{t\to\infty} E({}^wJ_t^{(0)}\,{}^wJ_t^{(2)\,T}) = \frac{\gamma^2\kappa^2(d+1)}{4(1-\kappa^2)} I + O(\gamma^2\mu),$$
$$\lim_{t\to\infty} E({}^wJ_t^{(1)}\,{}^wJ_t^{(1)\,T}) = \frac{\gamma^2(1+\kappa^2)(d+1)}{4(1-\kappa^2)} I + O(\gamma^2\mu).$$

yielding the following approximate expression for ${}^w\Gamma$,

$${}^w\bar{\Gamma} = \frac{\gamma^2}{2\mu\alpha} + \frac{\gamma^2}{4}\left(1 + (d+1)\frac{1+2\kappa^2}{1-\kappa^2}\right) + O(\gamma^2\mu). \tag{22}$$

The first order correction depends upon κ. This is illustrated in figures 5 and 5, where the asymptotic tracking error variance, defined as

$$\lim_{t\to\infty} \|\tilde{\theta}_t\|^2$$

is displayed as a function of the stepsize μ. In both cases, we set: $\gamma = 0.05$, $d = 10$, $\sigma_v^2 = 3$ and, for every value o κ, $\sigma_u^2 = 1 - \kappa^2$ (so that $\alpha = 1$). Two values of κ are considered: $\kappa = 0$ (figure 5) and $\kappa = 0.9$ (figure 5)). On the figures, the value of the asymptotic tracking error variance obtained by Monte-Carlo simulations (solid-line) are compared with the approximate expressions obtained by (i) retaining only the first term in (21) and (22) (dashed line) or (ii) including the two terms in (21) and (22) (dashed-dotted line). As seen on these figures, the autoregressive coefficient strongly affects the asymptotic tracking error variance, as predicted by the second-order approximation.

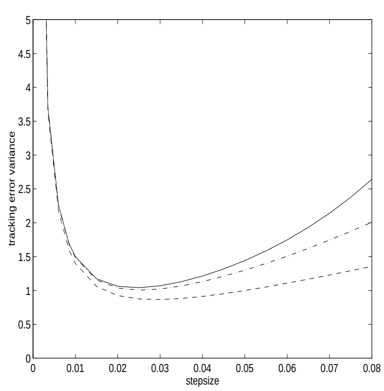

Figure 1: $\kappa = 0$. solid line: Monte-Carlo simulation. Dashed line: first order approximation. Dashed-dotted line: second-order approximation

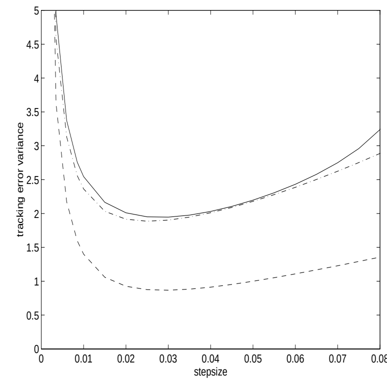

Figure 2: $\kappa = 0.9$. solid line: Monte-Carlo simulation. Dashed line: first order approximation. Dashed-dotted line: second-order approximation

6. REFERENCES

[1] P.Priouret and A. Veretennikov, A remark on the stability of the LMS. tracking algorithm, To appear in *Stochastic Analysis Appl.*, 1998

[2] L. Guo and L. Ljung and G.-J. Wang (1997), Necessary and Sufficient Conditions for Stability of LMS, *IEEE Trans. on Automatic Control*, vol. 42, N0 6, pp 761–769

[3] L. Guo and L. Ljung (1995), Exponential Stability of General Tracking Algoritms, *IEEE Trans. on Automatic Control*, vol.40 ,pp.1376-1387

[4] L.Guo, Stability of recursive stochastic tracking algorithms, *SIAM journal of control and Optimisation*, vol. 23, pp.1195-1225, 1994

[5] V.Solo and X.Kong, *Adaptive Signal Processing Algorithms: Stability and Performance, Prentice Hall , 1995*

[6] E.Eweda and O.Macchi., Tracking error bouds of adaptive nonstationary filtering, *Automatica*, vol. 2(3), pp. 293-302, 1985

[7] L. Ljung and P.Priouret (1991) A result on the mean square error obtained using general tracking algorithms. *Int. Jour. on Adaptative Control and Sig. Proc.*, vol. 5, n0 4, pp. 231–250

[8] P. Doukhan and L. Sana Weak dependence and moment inequalities Pre-print Université de Paris-Sud, 1997.