Tracking Performance Analysis of the Set-Membership NLMS Adaptive Filtering Algorithm

Reza Arablouei[*] and Kutluyıl Doğançay[†]
Institute for Telecommunications Research
[†]School of Electrical and Information Engineering
University of South Australia, Mawson Lakes 5095 SA, Australia
[*]E-mail: arary003@mymail.unisa.edu.au Tel: +61-8-830-23316

***Abstract*—In this paper, we analyze the tracking performance of the set-membership normalized least mean squares (SM-NLMS) adaptive filtering algorithm using the energy conservation argument. The analysis leads to a nonlinear equation whose solution gives the steady-state mean squared error (MSE) of the SM-NLMS algorithm in a nonstationary environment. We prove that there is always a unique positive solution for this equation. The results predicted by the analysis show good agreement with the simulation experiments.**

I. INTRODUCTION

Set-membership (SM) adaptive filtering algorithms are set-theoretic techniques, which estimate the members of a *set* of solutions that are consistent with observations and *a priori* signal or model constraints. Most SM adaptive filtering algorithms have been developed assuming a *bounded-error* constraint where the *a priori* knowledge emerges as a magnitude bound on the estimation error. The capacity of incorporating the *a priori* information in the estimation process has endowed the SM-based approaches with several advantages over their more classical counterparts in the realm of adaptive signal processing. Superior tracking capabilities and reduced average computational complexity are possibly the most notable ones. The former is achieved because every SM adaptive filtering algorithm executes a form of set-theoretic optimization that allows faster readjustment to the changes in the underlying system. The latter is also due to the inherent data-discerning update strategy of the SM approaches, which enables them to check for the innovation in the new data and perform an update only when it can improve the quality of the estimation. In the other words, the SM-based techniques are able to detect and discard the unhelpful data and save the expense of associated updating [1]-[5].

The set-membership normalized least mean squares (SM-NLMS) algorithm [6] is a well-known SM adaptive filtering algorithm that can be viewed as a SM equivalent of the stochastic-gradient NLMS algorithm. Formulation of the SM-NLMS algorithm is the same as that of the NLMS algorithm except having an optimized adaptive step-size instead of a fixed one. The SM-NLMS algorithm has attracted a lot of interest in adaptive signal processing mainly because of its relative simplicity, low computational complexity, robustness against noise, and numerical stability in finite precision implementations [7].

The SM adaptive filtering algorithms generally have high degrees of nonlinearities, which make their performance analysis complicated, especially at the transient state. The SM-NLMS algorithm involves both data nonlinearity [8] and error nonlinearity [9]. Few works have been published on the theoretical performance analysis of the SM adaptive filtering algorithms. Steady-state mean squared error (MSE) of the SM-NLMS algorithm in a stationary environment is analyzed in [10] and [11]. Steady-state MSE of the set-membership affine projection (SM-AP) algorithm in a stationary environment is also analyzed in [20] and [21]. Since the SM-NLMS algorithm is a special case of the SM-AP algorithm (when the projection order is equal to one), the results of [20] and [21] can apply to the SM-NLMS algorithm as well. To the best of our knowledge, the work of [22] is the only one to date that reports a tracking performance analysis of the SM-NLMS algorithm, i.e., steady-state MSE analysis in a nonstationary environment. This work originally deals with the SM-AP algorithm. It is based on an empirical formula for the probability of updating and employs some assumptions, which are in general difficult to justify.

In this paper, we analyze the tracking performance of the SM-NLMS algorithm with a more rigorous approach than the one taken in [22]. Following the customary trend in the literature, we take the steady-state MSE in a nonstationary environment as the measure of tracking performance and assume that the target system varies in time according to the random walk model. The classic approach is to analyze the transient behavior of the adaptive filter and then obtain the steady-state MSE as a limiting behavior of the transient MSE [15]. However, due to high nonlinearity of the SM-NLMS algorithm, this approach would be very difficult. To circumvent the analysis of the transient behavior and its complications, we initiate the steady-state analysis from the energy conservation relation [12]-[14] of the SM-NLMS algorithm. The outcome of the analysis is a nonlinear equation whose solution yields the theoretical steady-state MSE of the SM-NLMS algorithm in a nonstationary environment. Simulations confirm the accuracy of the theoretical findings.

In section 2, we describe the SM-NLMS algorithm. We analyze its tracking performance in section 3, provide some

This work was supported in part by a *Commonwealth Scientific and Industrial Research Organisation (CSIRO)* scholarship.

simulation results in section 4, and conclude the paper in section 5.

II. The Set-Membership NLMS Algorithm

Let us consider the affine-in-parameter model of the form

$$d(n) = \mathbf{w}_o^T(n)\mathbf{x}(n) + \eta(n) \quad (1)$$

where $d(n) \in \mathbb{R}$ is the reference signal at time index $n \in \mathbb{N}$, $\mathbf{w}_o(n) \in \mathbb{R}^L$ is the column vector of the underlying system parameters, $\mathbf{x}(n) \in \mathbb{R}^L$ is the input vector, $\eta(n) \in \mathbb{R}$ accounts for background noise, $\mathbb{N}$ and $\mathbb{R}$ respectively denote the sets of all nonnegative integers and all real numbers, and superscript T stands for transposition. Moreover, let us assume that the underlying system varies in time according to the well-known random walk model [14], i.e.,

$$\mathbf{w}_o(n) = \mathbf{w}_o(n-1) + \mathbf{q}(n) \quad (2)$$

where $\mathbf{q}(n) \in \mathbb{R}^L$ denotes the random perturbation.

SM adaptive filtering algorithms estimate a set of feasible solutions rather than a single-point estimate by specifying a bound on the magnitude of the estimation error over the model space of interest, $\mathcal{S}$. Hence, the *feasibility set* that contains all the parameter vectors resulting in an error less than a predetermined bound, $\gamma > 0$, is defined as

$$\Theta = \bigcap_{(\mathbf{x},d)\in\mathcal{S}} \{\mathbf{w} \in \mathbb{R}^L : |d - \mathbf{w}^*\mathbf{x}| \le \gamma\}.$$

In general, SM adaptive filtering algorithms aim to approximate adaptively the minimal-set estimate of Θ that is called the *membership set* and is defined as

$$\Psi(n) = \bigcap_{i=1}^{n} \mathcal{H}(i)$$

where $\mathcal{H}(n)$ is the observation-induced *constraint set* at time instant n:

$$\mathcal{H}(n) = \{\mathbf{w} \in \mathbb{R}^L : |d(n) - \mathbf{w}^*\mathbf{x}(n)| \le \gamma\}.$$

One way is to outer-approximate $\Psi(n)$ at each time instant by a minimal-volume spheroid described by

$$S(n) = \{\mathbf{w} \in \mathbb{R}^L : \|\mathbf{w} - \mathbf{w}(n)\|^2 \le \sigma^2(n)\}$$

where $\|\cdot\|$ denotes the Euclidean norm, $\mathbf{w}(n)$ is the center of the spheroid (which is usually taken as the point estimate), and $\sigma(n)$ is its radius. This is the basic idea behind the SM-NLMS algorithm, which, at each iteration, computes $S(n)$ as the smallest spheroid that contains the intersection of $S(n-1)$ and $\mathcal{H}(n)$. It is shown in [6] that this can be realized using the following recursions:

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mu(n)\frac{\mathbf{x}(n)}{\|\mathbf{x}(n)\|^2}e(n), \quad (3)$$

$$\sigma^2(n) = \sigma^2(n-1) - \mu^2(n)\frac{e^2(n)}{\|\mathbf{x}(n)\|^2}$$

where the estimation error, $e(n)$, and the variable step-size, $\mu(n)$, are given by

$$e(n) = d(n) - \mathbf{w}^T(n-1)\mathbf{x}(n) \quad (4)$$

and

$$\mu(n) = \begin{cases} 1 - \frac{\gamma}{|e(n)|} & \text{if } |e(n)| > \gamma \\ 0 & \text{if } |e(n)| \le \gamma \end{cases}. \quad (5)$$

III. Tracking Performance Analysis

Substituting (5) into (3) yields

$$\mathbf{w}(n) = \begin{cases} \mathbf{w}(n-1) + \left(1 - \frac{\gamma}{|e(n)|}\right)\frac{\mathbf{x}(n)}{\|\mathbf{x}(n)\|^2}e(n) & \text{if } |e(n)| > \gamma \\ \mathbf{w}(n-1) & \text{if } |e(n)| \le \gamma \end{cases}. \quad (6)$$

Defining

$$f(n) = \begin{cases} e(n) - \gamma & \text{if } e(n) > \gamma \\ 0 & \text{if } -\gamma \le e(n) \le \gamma, \\ e(n) + \gamma & \text{if } e(n) < -\gamma \end{cases} \quad (7)$$

we can rewrite (6) as

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \frac{\mathbf{x}(n)}{\|\mathbf{x}(n)\|^2}f(n). \quad (8)$$

Subtracting both sides of (8) from $\mathbf{w}_o(n)$ together with using (2) results in

$$\mathbf{v}(n) = \mathbf{v}(n-1) - \frac{\mathbf{x}(n)}{\|\mathbf{x}(n)\|^2}f(n) + \mathbf{q}(n) \quad (9)$$

where $\mathbf{v}(n) = \mathbf{w}_o(n) - \mathbf{w}(n)$ is the weight-error vector.

The noiseless *a priori* and *a posteriori* estimation errors are respectively defined as

$$e_a(n) = \mathbf{v}(n-1)^T\mathbf{x}(n) \quad (10)$$

and

$$e_p(n) = [\mathbf{v}(n) - \mathbf{q}(n)]^T\mathbf{x}(n).$$

From (1), (4), and (10), we have

$$e(n) = e_a(n) + \eta(n). \quad (11)$$

Multiplying the transpose of both sides of (9) by $\mathbf{x}(n)$ from the right results in

$$e_p(n) = e_a(n) - f(n). \quad (12)$$

By substituting $f(n) = e_a(n) - e_p(n)$ into (9), we get

$$\mathbf{v}(n) - \mathbf{q}(n) + \frac{\mathbf{x}(n)}{\|\mathbf{x}(n)\|^2}e_a(n) = \mathbf{v}(n-1) + \frac{\mathbf{x}(n)}{\|\mathbf{x}(n)\|^2}e_p(n). \quad (13)$$

We obtain the following energy conservation relation by evaluating the energies (i.e., the squared Euclidean norms) of both sides of (13):

$$\|\mathbf{v}(n) - \mathbf{q}(n)\|^2 + \frac{e_a^2(n)}{\|\mathbf{x}(n)\|^2} = \|\mathbf{v}(n-1)\|^2 + \frac{e_p^2(n)}{\|\mathbf{x}(n)\|^2}. \quad (14)$$

Taking the expectations of both sides of (14) gives

$$E[\|\mathbf{v}(n) - \mathbf{q}(n)\|^2] + E\left[\frac{e_a^2(n)}{\|\mathbf{x}(n)\|^2}\right] = E[\|\mathbf{v}(n-1)\|^2] + E\left[\frac{e_p^2(n)}{\|\mathbf{x}(n)\|^2}\right]. \quad (15)$$

In order to make the analysis tractable, let us adopt the following assumptions:

A0: The stochastic process $\|\mathbf{x}(n)\|^{-2}$ is strictly stationary and the initial input, $\mathbf{x}(0)$, is a nonzero vector that satisfies $E[\|\mathbf{x}(0)\|^{-2}] < \infty$.

A1: At the steady state, $\|\mathbf{x}(n)\|^{-2}$ is statistically independent of both $e_a^2(n)$ and $e_p^2(n)$.

A2: At the steady state, the noiseless *a priori* estimation error, $e_a(n)$, is white Gaussian with zero mean and variance of $\sigma_a^2 = E[e_a^2(n)]$.

A3: The background noise, $\eta(n)$, is white Gaussian with zero mean and variance of $\sigma_\eta^2 = E[\eta^2(n)]$. It is also independent of the input data.

A4: The perturbation $\mathbf{q}(n)$ is independent and identically distributed (i.i.d.) with zero mean and covariance matrix of $\mathbf{Q} = E[\mathbf{q}(n)\mathbf{q}^T(n)]$. It is also independent of the input data and the noise.

Under A4, we have

$$\begin{aligned} &E[\|\mathbf{v}(n) - \mathbf{q}(n)\|^2] \\ &\quad = E[\|\mathbf{v}(n)\|^2] - 2E[\mathbf{q}^T(n)\mathbf{v}(n)] + E[\|\mathbf{q}(n)\|^2] \\ &\quad = E[\|\mathbf{v}(n)\|^2] - 2E[\|\mathbf{q}(n)\|^2] + E[\|\mathbf{q}(n)\|^2] \\ &\quad = E[\|\mathbf{v}(n)\|^2] - \mathrm{Tr}(\mathbf{Q}) \end{aligned} \quad (16)$$

where $\mathrm{Tr}(\cdot)$ denotes the matrix trace.

We know that after convergence, gradient of the mean square deviation approaches zero [6], i.e.,

$$\lim_{n\to\infty} E[\|\mathbf{v}(n-1)\|^2] - E[\|\mathbf{v}(n)\|^2] \to 0. \quad (17)$$

Using (16) and (17), at the steady state, (15) becomes

$$E\left[\frac{e_a^2(n)}{\|\mathbf{x}(n)\|^2}\right] = E\left[\frac{e_p^2(n)}{\|\mathbf{x}(n)\|^2}\right] + \mathrm{Tr}(\mathbf{Q}) \quad (18)$$

and using A0 and A1, (18) simplifies to

$$E[e_a^2(n)] = E\left[e_p^2(n)\right] + \frac{\mathrm{Tr}(\mathbf{Q})}{E[\|\mathbf{x}(n)\|^{-2}]}. \quad (19)$$

Substituting (12) into (19), we find that at the steady state

$$E[f^2(n) - 2e_a(n)f(n)] + \frac{\mathrm{Tr}(\mathbf{Q})}{E[\|\mathbf{x}(n)\|^{-2}]} = 0. \quad (20)$$

Considering (11), A2 and A3 imply that, at the steady state, the *a priori* estimation error, $e(n)$, is also zero-mean white Gaussian. In addition, the variance of $e(n)$ at the steady-state is computed as

$$\sigma^2 = E[e^2(n)] = \sigma_a^2 + \sigma_\eta^2. \quad (21)$$

Based on this fact and using (7), we can write

$$\begin{aligned} &E[f^2(n)] \\ &= \int_{-\infty}^{-\gamma}(x+\gamma)^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}dx + \int_{\gamma}^{\infty}(x-\gamma)^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}dx \\ &\quad = (\sigma^2+\gamma^2)\left[1 - \mathrm{erf}\left(\frac{\gamma}{\sqrt{2\sigma^2}}\right)\right] - \gamma\sqrt{\frac{2\sigma^2}{\pi}}e^{-\frac{\gamma^2}{2\sigma^2}} \end{aligned} \quad (22)$$

where $\mathrm{erf}(\cdot)$ is the error function and is defined as

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt.$$

We can also verify that, at the steady state, $e_a(n)$ and $e(n)$ are jointly Gaussian and the random variable

$$\zeta = e_a(n) - \frac{\sigma_a^2}{\sigma^2}e(n) \quad (23)$$

is independent of $e(n)$ and thus $f(n)$. Consequently, using (21) and (23), we can show that

$$\begin{aligned} E[e_a(n)f(n)] &= E[\zeta f(n)] + \frac{\sigma^2-\sigma_\eta^2}{\sigma^2}E[e(n)f(n)] \\ &= E[\zeta]E[f(n)] + \frac{\sigma^2-\sigma_\eta^2}{\sigma^2}E[e(n)f(n)] \\ &= \frac{\sigma^2-\sigma_\eta^2}{\sigma^2}E[e(n)f(n)]. \end{aligned} \quad (24)$$

Besides, we have

$$\begin{aligned} &E[e(n)f(n)] \\ &= \int_{-\infty}^{-\gamma} x(x+\gamma)\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{x^2}{2\sigma^2}}dx + \int_{\gamma}^{\infty} x(x-\gamma)\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{x^2}{2\sigma^2}}dx \\ &\quad = \sigma^2\left[1 - \mathrm{erf}\left(\frac{\gamma}{\sqrt{2\sigma^2}}\right)\right]. \end{aligned} \quad (25)$$

Hence, by substituting (25) into (24), we get

$$E[e_a(n)f(n)] = \left(\sigma^2 - \sigma_\eta^2\right)\left[1 - \mathrm{erf}\left(\frac{\gamma}{\sqrt{2\sigma^2}}\right)\right]. \quad (26)$$

Substitution of (22) and (26) into (20) together with using A0 leads to the following nonlinear equation:

$$\begin{aligned} &\left(-\sigma^2 + \gamma^2 + 2\sigma_\eta^2\right)\left[1 - \mathrm{erf}\left(\frac{\gamma}{\sqrt{2\sigma^2}}\right)\right] - \gamma\sqrt{\frac{2\sigma^2}{\pi}}e^{-\frac{\gamma^2}{2\sigma^2}} \\ &\quad + \frac{\mathrm{Tr}(\mathbf{Q})}{E[\|\mathbf{x}(0)\|^{-2}]} = 0. \end{aligned} \quad (27)$$

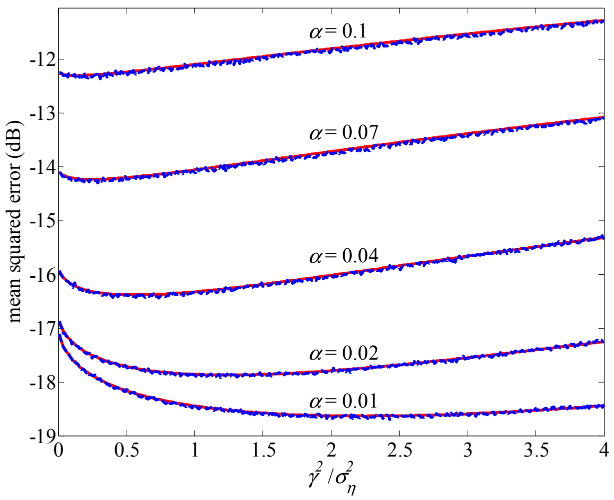
The unique positive solution of (27) for σ^2 yields the steady-state MSE of the SM-NLMS algorithm when the underlying system is time-varying. In the Appendix, we prove existence and uniqueness of the solution. It is clear that (27) is the extension of the equation given in [11] (for the unrelaxed and unregularized SM-NLMS algorithm) to the case of nonstationary environment.

IV. Simulations

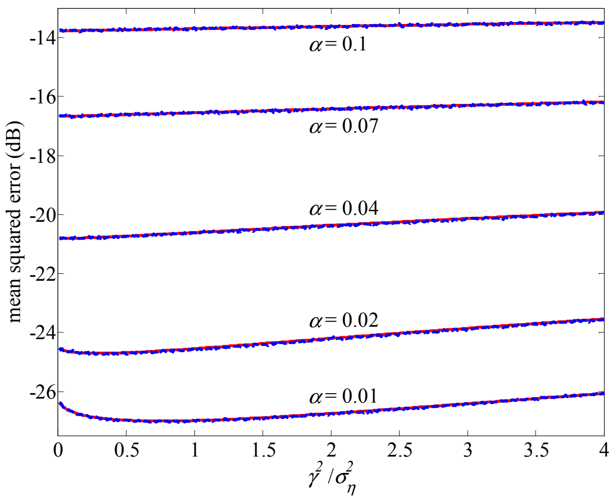
We consider estimation and tracking of a flat-fading multiple-input single-output (MISO) wireless channel with $N_t = 4$ transmitter antennas. Sub-channels between each transmitter and the receiver antennas fade independently according to the random walk model. Thus, the channel vector, $\mathbf{h}(n) \in \mathbb{R}^{N_t}$, is modeled via

$$\mathbf{h}(n) = \mathbf{h}(n-1) + \alpha\mathbf{k}(n)$$

where α is a parameter that controls the rate of the variations and $\mathbf{k}(n) \in \mathbb{R}^{N_t}$ is a circular Gaussian random vector with zero mean and identity covariance matrix. The channel estimator is an adaptive filter with $L = 4$ taps. The transmitted symbols (input to the channel estimator) are i.i.d. and modulated using the BPSK scheme with a modulus of $1/\sqrt{N_t} = 1/2$.



(a)



(b)

Fig. 1. Comparing experimental (dashed lines) and theoretical (solid lines) MSEs of the SM-NLMS algorithm in nonstationary environments with different values of α when $\sigma_\eta^2 = 0.01$ (a) and $\sigma_\eta^2 = 0.001$ (b).

In Fig. 1, we depict the experimental and theoretical steady-state MSEs as a function of γ^2/σ_η^2 for different values of α and two values of the power of the additive white Gaussian noise, i.e., $\sigma_\eta^2 = 0.01$ and 0.001. The experimental steady-state MSEs are obtained by averaging 100 steady-state values and ensemble-averaging over 10^3 independent runs. The theoretical steady-state MSEs are calculated by solving (27) where the last term of its left-hand side simplifies to

$$\frac{\mathrm{Tr}(E[\alpha\mathbf{k}(n)\alpha\mathbf{k}^T(n)])}{E[\|\mathbf{x}(0)\|^{-2}]} = N_t\alpha^2.$$

Fig. 1 testifies that our theoretical results are in excellent match with the simulations. It is also noteworthy that the theoretical predictions of [22] severely mismatch the experimental results in the examined scenario. The mismatch is more pronounced for higher values of α and γ. To avoid confusion and maintain legibility of the figures, we do not include the results of [22] in Fig. 1.

V. Concluding Remarks

Assumptions: A0 is required for tractability of the analysis. However, assuming a stationary $\|\mathbf{x}(n)\|^{-2}$ is not as strict as assuming a stationary input. Therefore, our analysis is valid for any input signal as long as $\|\mathbf{x}(n)\|^{-2}$ is stationary. Despite seeming unnatural, A1 is reasonably realistic and simplifies the analysis in a great deal. Several similar assumptions have been made in the literature (see, e.g., [11], [13], [14], [15]). Without utilizing A2, calculation of the expectations in (20) would be arduous. This assumption is commonly made to deal with error nonlinearities (see, e.g., [9], [11]). It is justified for long adaptive filters via central limit arguments [16]. Lastly, A3 is natural and regularly used while A4 is typical in the context of tracking performance analysis of the adaptive filters (see, e.g., [13], [17]).

Simulations: The primary objective of the provided numerical results is to corroborate our theoretical findings. That is why the simulated MISO wireless channel varies in time according to the random walk model, which is the assumed model in the analysis. Admittedly, this is not the most appropriate way to model variations of a wireless channel. However, it is well established that a Rayleigh-fading channel can be suitably modeled by a first-order autoregressive model [18]. For slow variations, the corresponding autoregressive model can be approximated by a random walk model [14].

Summary: Tracking performance of the SM-NLMS algorithm was analyzed utilizing the energy conservation argument. The analysis resulted in a nonlinear equation that by solving it, the theoretical steady-state MSE of the SM-NLMS algorithm in a nonstationary system is calculated. Existence and uniqueness of the solution of this equation was also proved. Simulations showed that the theoretical MSE values match the experimental ones well.

Appendix

Introducing the change of variable $x = \gamma/\sqrt{2\sigma^2}$ and defining $a = 1 + 2\sigma_\eta^2/\gamma^2$ and $b = \mathrm{Tr}(\mathbf{Q})/E[\|\mathbf{x}(0)\|^{-2}]$, we can rewrite (27) as

$$ax^2 - \frac{1}{2} + \frac{b}{\gamma^2}\frac{x^2}{1-\mathrm{erf}(x)} = \frac{1}{\sqrt{\pi}}\frac{xe^{-x^2}}{1-\mathrm{erf}(x)}. \quad (28)$$

Lemma 1: Let

$$Y(x) = ax^2 - \frac{1}{2} + \frac{b}{\gamma^2}\frac{x^2}{1-\mathrm{erf}(x)} \quad \text{for } x > 0. \quad (29)$$

Then, the following holds:

1) $\lim_{x\to 0} Y(x) = -1/2$;
2) $\lim_{x\to\infty} Y(x) = \infty$;
3) $Y(x)$ is monotone increasing.

Proof:

1) This is obvious.
2) This can be easily verified since $a \geq 1$, $b \geq 0$, $1 - \mathrm{erf}(x) > 0$ for $x > 0$, and $\lim_{x\to\infty}[1 - \mathrm{erf}(x)] = 0$.

3) This is also easy to verify since, for $x > 0$, $ax^2 - 1/2$ is monotone increasing and $1 - \mathrm{erf}(x)$ is monotone decreasing in $(0,1]$ so $\frac{b}{\gamma^2}\frac{x^2}{1-\mathrm{erf}(x)}$ is monotone increasing. ■

Lemma 2: Let

$$Z(x) = \frac{1}{\sqrt{\pi}}\frac{xe^{-x^2}}{1-\mathrm{erf}(x)} \quad \text{for } x > 0. \tag{30}$$

Then, the following holds:

1) $\lim_{x\to 0} Z(x) = 0$;
2) $\lim_{x\to\infty} Z(x) = \infty$;
3) $Z(x)$ is monotone increasing.

Proof:

1) This is obvious.
2) Using L'Hôpital's rule, we have

$$\begin{aligned}\lim_{x\to\infty} Z(x) &= \frac{1}{\sqrt{\pi}}\lim_{x\to\infty}\frac{\frac{d}{dx}\left[xe^{-x^2}\right]}{\frac{d}{dx}[1-\mathrm{erf}(x)]}\\ &= \frac{1}{\sqrt{\pi}}\lim_{x\to\infty}\frac{-(2x^2-1)e^{-x^2}}{-\frac{2}{\sqrt{\pi}}e^{-x^2}}\\ &= \infty.\end{aligned}$$

3) Let us define

$$\Omega(x) = \frac{e^{-x^2}}{1-\mathrm{erf}(x)}$$

and calculate its derivative

$$\frac{d\Omega(x)}{dx} = \frac{\frac{2}{\sqrt{\pi}} - 2xe^{x^2}[1-\mathrm{erf}(x)]}{e^{2x^2}[1-\mathrm{erf}(x)]^2}.$$

Invoking Gordon's inequality [19], i.e.,

$$\sqrt{\pi}e^{x^2}[1-\mathrm{erf}(x)] < \frac{1}{x} \quad \text{for } x > 0,$$

we can show that $d\Omega(x)/dx > 0$ for $x > 0$. This means $\Omega(x)$ and subsequently $Z(x) = x/\sqrt{\pi}\,\Omega(x)$ are monotone increasing for $x > 0$. ■

Lemma 3: Given $Y(x)$ and $Z(x)$ defined by (29) and (30), the following holds:

$$\lim_{x\to\infty}\frac{Y(x)}{Z(x)} = \begin{cases}\infty & \text{if } b > 0\\ a & \text{if } b = 0\end{cases}.$$

Proof: We can write

$$\begin{aligned}\lim_{x\to\infty}\frac{Y(x)}{Z(x)} &= \lim_{x\to\infty}\frac{\left(ax^2-\frac{1}{2}\right)[1-\mathrm{erf}(x)]+\frac{b}{\gamma^2}x^2}{\frac{1}{\sqrt{\pi}}xe^{-x^2}}\\ &= \sqrt{\pi}\left(a\lim_{x\to\infty}\frac{x[1-\mathrm{erf}(x)]}{e^{-x^2}} - \frac{1}{2}\lim_{x\to\infty}\frac{1-\mathrm{erf}(x)}{xe^{-x^2}}\right.\\ &\quad\left.+\frac{b}{\gamma^2}\lim_{x\to\infty} xe^{x^2}\right).\end{aligned} \tag{31}$$

Using L'Hôpital's rule, we have

$$\begin{aligned}\lim_{x\to\infty}\frac{x[1-\mathrm{erf}(x)]}{e^{-x^2}} &= \lim_{x\to\infty}\frac{\frac{d}{dx}[x-\mathrm{erf}(x)x]}{\frac{d}{dx}\left[e^{-x^2}\right]}\\ &= \lim_{x\to\infty}\frac{1-\mathrm{erf}(x)-\frac{2}{\sqrt{\pi}}xe^{-x^2}}{-2xe^{-x^2}}\\ &= \frac{1}{\sqrt{\pi}}\end{aligned} \tag{32}$$

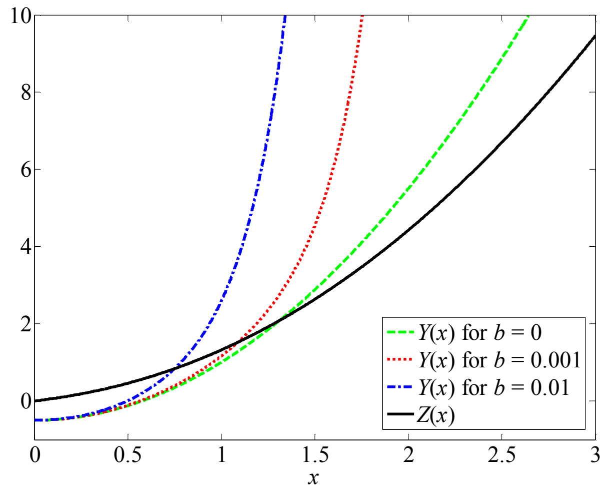


Fig. 2. The functions $Z(x)$ and $Y(x)$ for different values of b.

and

$$\begin{aligned}\lim_{x\to\infty}\frac{1-\mathrm{erf}(x)}{xe^{-x^2}} &= \lim_{x\to\infty}\frac{\frac{d}{dx}[1-\mathrm{erf}(x)]}{\frac{d}{dx}\left[xe^{-x^2}\right]}\\ &= \lim_{x\to\infty}\frac{-\frac{2}{\sqrt{\pi}}e^{-x^2}}{-(2x^2-1)e^{-x^2}}\\ &= 0.\end{aligned} \tag{33}$$

Substituting (32) and (33) into (31) results in

$$\begin{aligned}\mathrm{Lim}_{x\to\infty}\frac{Y(x)}{Z(x)} &= a + \sqrt{\pi}\frac{b}{\gamma^2}\lim_{x\to\infty} xe^{x^2}\\ &= \begin{cases}\infty & \text{if } b > 0\\ a & \text{if } b = 0\end{cases}.\end{aligned}$$ ■

From Lemmas 1-3 and the fact that $a \geq 1$, one can deduce that the equation $Y(x) = Z(x)$, i.e., (28), surely has a unique positive solution.

In Fig. 2, we plot $Y(x)$ and $Z(x)$ for $\sigma_\eta^2 = 0.01$, $\gamma^2/\sigma_\eta^2 = 4$, and different values of b. It is observed that a larger b (higher degree of nonstationarity) leads to a lower $x = \gamma/\sqrt{2\sigma^2}$ and consequently a higher σ^2.

REFERENCES

[1] M. Milanese, J. P. Norton, H. Piet-Lahanier, and E. Walter (eds.), *Bounding Approaches to System Identification,* London: Plenum, 1996.

[2] J. R. Deller, Jr. and Y. F. Huang, "Set-membership identification and filtering for signal processing applications," *Circuits Systems Signal Process.*, vol. 21, no. 1, pp. 69–82, 2002.

[3] E.-W. Bai, K. M. Nagpal, and R. Tempo, "Bounded-error estimation: noise models and recursive algorithms," *Automatica*, vol. 32, pp. 1–15, 2002.

[4] J. R. Deller, Jr., M. Nayeri, and S. F. Odeh, "Least-square identification with error bounds for real-time signal processing and control," *Proc. IEEE*, vol. 81, pp. 813–849, June 1993.

[5] P. L. Combettes, "The foundations of set-theoretic estimation," *Proc. IEEE*, vol. 81, pp. 182–208, 1993.

[6] S. Gollamudi, S. Nagaraj, S. Kapoor, and Y. F. Huang, "Set-membership filtering and a set-membership normalized LMS

algorithm with an adaptive step size," *IEEE Signal Process. Lett.*, vol. 5, no. 5, pp. 111–114, May 1998.

[7] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*, 3rd ed. New York: Springer, 2008.

[8] T. Y. Al-Naffouri and A. H. Sayed, "Transient analysis of data-normalized adaptive filters," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 639–652, Mar. 2003.

[9] T. Y. Al-Naffouri and A. H. Sayed, "Transient analysis of adaptive filters with error nonlinearities," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 653–663, Mar. 2003.

[10] N. Takahashi and I. Yamada, "Steady-state performance of hyperslab projection algorithm," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Las Vegas, USA, April 2008, pp. 3829–3832.

[11] N. Takahashi and I. Yamada, "Steady-state mean-square performance analysis of a relaxed set-membership NLMS algorithm by the energy conservation argument," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3361-3372, 2009.

[12] M. Rupp and A. H. Sayed, "A time-domain feedback analysis of filtered-error adaptive gradient algorithms," *IEEE Trans. Signal Process.*, vol. 44, pp. 1428–1439, June 1996.

[13] N. R. Yousef and A. H. Sayed, "A unified approach to the steady-state and tracking analyses of adaptive filters," *IEEE Trans. Signal Process.*, vol. 49, no. 2, pp. 314–324, Feb. 2001.

[14] A. H. Sayed, *Adaptive Filters*, Hoboken, NJ: Wiley, 2008.

[15] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, NJ: Prentice-Hall, 2002.

[16] D. L. Duttweiler, "Adaptive filter performance with nonlinearities in the correlation multiplier," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 30, pp. 578–586, Aug. 1982.

[17] E. Eweda, "Comparison of RLS, LMS, and sign algorithms for tracking randomly time-varying channels," *IEEE Trans. Signal Process.*, vol. 42, pp. 2937-2944, Nov. 1994.

[18] C. Komninakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, "Multi-input multi-output fading channel tracking and equalization using Kalman estimation," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1065–1076, May 2002.

[19] R. D. Gordon, "Values of Mills' ratio of area bounding ordinate and of the normal probability integral for large values of the argument," *Ann. Math. Statistics*, vol. 12, pp. 364–366, 1941.

[20] M. V. S. Lima and P .S. R. Diniz, "Steady-state analysis of the set-membership affine projection algorithm," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Dallas, USA, March 2010, pp. 3802–3805.

[21] P. S. R. Diniz, "Convergence performance of the simplified set-membership affine projection algorithm," *Circuits Systems Signal Process.*, vol. 30, pp. 439-462, 2011.

[22] P. S. R. Diniz, "Analysis of a set-membership affine projection algorithm in nonstationary environment," in *Proc. European Signal Process. Conf.*, Glasgow, Scotland, Aug. 2009, pp. 2623–2627.