CHAPTER 16

Performance of LMS

We now move on to illustrate the use of the variance relation (15.40) in evaluating the steady-state performance of adaptive filters. We start with the simplest of algorithms, namely, LMS.

16.1 VARIANCE RELATION

Thus, assume that the data $\{\boldsymbol{d}(i), \boldsymbol{u}_i\}$ satisfy model (15.16) and consider the LMS recursion

$$\boldsymbol{w}_i = \boldsymbol{w}_{i-1} + \mu \boldsymbol{u}_i^* \boldsymbol{e}(i) \tag{16.1}$$

for which

$$g[\boldsymbol{e}(i)] = \boldsymbol{e}(i) = \boldsymbol{e}_a(i) + \boldsymbol{v}(i) \tag{16.2}$$

Relation (15.40) then becomes

$$\mu \mathsf{E}\,\|\boldsymbol{u}_i\|^2 |\boldsymbol{e}_a(i) + \boldsymbol{v}(i)|^2 = 2\mathsf{Re}\big(\mathsf{E}\,\boldsymbol{e}_a^*(i)[\boldsymbol{e}_a(i) + \boldsymbol{v}(i)]\big), \quad i \to \infty \tag{16.3}$$

Several terms in this equality get cancelled. We shall carry out the calculations rather slowly in this section for illustration purposes only. Later, when similar calculations are called upon, we shall be less detailed.

To begin with, the expression on the left-hand side of (16.3) expands to

$$\begin{aligned} \mu \mathsf{E}\,\|\boldsymbol{u}_i\|^2 |\boldsymbol{e}_a(i) + \boldsymbol{v}(i)|^2 &= \mu \mathsf{E}\,\|\boldsymbol{u}_i\|^2 \big[\, |\boldsymbol{e}_a(i)|^2 + |\boldsymbol{v}(i)|^2 + \boldsymbol{e}_a^*(i)\boldsymbol{v}(i) + \boldsymbol{e}_a(i)\boldsymbol{v}^*(i)\,\big] \\ &= \mu \mathsf{E}\,\|\boldsymbol{u}_i\|^2 |\boldsymbol{e}_a(i)|^2 \;+\; \mu\sigma_v^2 \mathsf{E}\,\|\boldsymbol{u}_i\|^2 \\ &= \mu \mathsf{E}\,\|\boldsymbol{u}_i\|^2 |\boldsymbol{e}_a(i)|^2 \;+\; \mu\sigma_v^2 \mathsf{Tr}(R_u) \end{aligned} \tag{16.4}$$

where we used the fact that $\boldsymbol{v}(i)$ is independent of both $\boldsymbol{u}_i$ and $\boldsymbol{e}_a(i)$ (recall Lemma 15.1), so that the cross-terms involving $\{\boldsymbol{v}(i), \boldsymbol{e}_a(i), \boldsymbol{u}_i\}$ cancel out. We also used the fact that

$$\mathsf{E}\,\|\boldsymbol{u}_i\|^2 = \mathsf{Tr}(R_u) \quad \text{and} \quad \mathsf{E}\,|\boldsymbol{v}(i)|^2 = \sigma_v^2$$

Similarly, the expression on the right-hand side of (16.3) simplifies to $2\mathsf{E}\,|\boldsymbol{e}_a(i)|^2$, which is simply $2\zeta^{\mathsf{LMS}}$ as $i \to \infty$. Therefore, equality (16.3) amounts to

$$\boxed{\zeta^{\mathsf{LMS}} = \frac{\mu}{2}\left[\mathsf{E}\,\|\boldsymbol{u}_i\|^2 |\boldsymbol{e}_a(i)|^2 \;+\; \sigma_v^2 \mathsf{Tr}(R_u)\right], \quad \text{as } i \to \infty} \tag{16.5}$$

This expression has been arrived at without approximations. Still, it requires that we evaluate the steady-state value of the expectation $\mathsf{E}\,\|\boldsymbol{u}_i\|^2 |\boldsymbol{e}_a(i)|^2$ in order to arrive at the EMSE

Adaptive Filters, by Ali H. Sayed

of LMS. Some assumptions will now need to be introduced in order to proceed with the analysis, even for this simplest of algorithms!

We shall examine three scenarios. One scenario assumes sufficiently small step-sizes, while another relies on a useful separation assumption. The third scenario assumes regressors with Gaussian distribution.

16.2 SMALL STEP-SIZES

Expression (16.5) suggests that small step-sizes lead to small $\mathsf{E}\,|e_a(i)|^2$ in steady-state and, consequently, to a high likelihood of small values for $e_a(i)$ itself. So assume μ is small enough so that, in *steady-state*, the contribution of the term $\mathsf{E}\,\|\boldsymbol{u}_i\|^2|\boldsymbol{e}_a(i)|^2$ can be neglected, say,

$$\mathsf{E}\,\|\boldsymbol{u}_i\|^2|\boldsymbol{e}_a(i)|^2 \;\ll\; \sigma_v^2\mathsf{Tr}(R_u)$$

Then, we find from (16.5) that the EMSE can be approximated by

$$\boxed{\zeta^{\mathsf{LMS}} = \frac{\mu\sigma_v^2\mathsf{Tr}(R_u)}{2}} \qquad \text{(for sufficiently small } \mu) \tag{16.6}$$

16.3 SEPARATION PRINCIPLE

If the step-size is not sufficiently small, but still small enough to guarantee filter convergence — as will be discussed in Chapter 24, we can derive an alternative approximation for the EMSE from (16.5); the resulting expression will hold over a wider range of step-sizes. To do so, here and in several other places in this chapter and in subsequent chapters, we shall rely on the following assumption:

$$\boxed{\text{At steady-state, } \;\|\boldsymbol{u}_i\|^2\; \text{ is independent of } \;\boldsymbol{e}_a(i)} \tag{16.7}$$

We shall refer to this condition as the *separation* assumption or the separation principle. Alternatively, we could assume instead that

$$\boxed{\text{At steady-state, } \;\|\boldsymbol{u}_i\|^2\; \text{ is independent of } \;\boldsymbol{e}(i)} \tag{16.8}$$

with $\boldsymbol{e}_a(i)$ replaced by $\boldsymbol{e}(i)$. This condition is equivalent to (16.7) since $\boldsymbol{e}(i) = \boldsymbol{e}_a(i)+\boldsymbol{v}(i)$ and $\|\boldsymbol{u}_i\|^2$ is independent of $\boldsymbol{v}(i)$ (as follows from Lemma 15.1).

Of course, assumption (16.7) is only exact in some special cases, e.g., when the successive regressors have constant Euclidean norms, since then $\|\boldsymbol{u}_i\|^2$ becomes a constant; this situation occurs when the entries of $\boldsymbol{u}_i$ arise from a finite alphabet with constant magnitude — see Prob. IV.2. More generally, the assumption is reasonable at *steady-state* since the behavior of $\boldsymbol{e}_a(i)$ in the limit is likely to be less sensitive to the regression (input) data.

Assumption (16.7) allows us to separate the expectation $\mathsf{E}\,\|\boldsymbol{u}_i\|^2|\boldsymbol{e}_a(i)|^2$, which appears in (16.5), into the product of two expectations:

$$\mathsf{E}\left(\|\boldsymbol{u}_i\|^2\cdot|\boldsymbol{e}_a(i)|^2\right) \;=\; \left(\mathsf{E}\,\|\boldsymbol{u}_i\|^2\right)\cdot\left(\mathsf{E}\,|\boldsymbol{e}_a(i)|^2\right) \;=\; \mathsf{Tr}(R_u)\zeta^{\mathsf{LMS}}, \quad i\to\infty \tag{16.9}$$

In order to illustrate this approximation, we show in Fig. 16.1 the result of simulating a 20-tap LMS filter over 1000 experiments. The figure shows the ensemble-averaged curves that correspond to the quantities

$$\mathsf{E}\left(\|\boldsymbol{u}_i\|^2\cdot|\boldsymbol{e}_a(i)|^2\right) \qquad \text{and} \qquad \left(\mathsf{E}\,\|\boldsymbol{u}_i\|^2\right)\cdot\left(\mathsf{E}\,|\boldsymbol{e}_a(i)|^2\right)$$

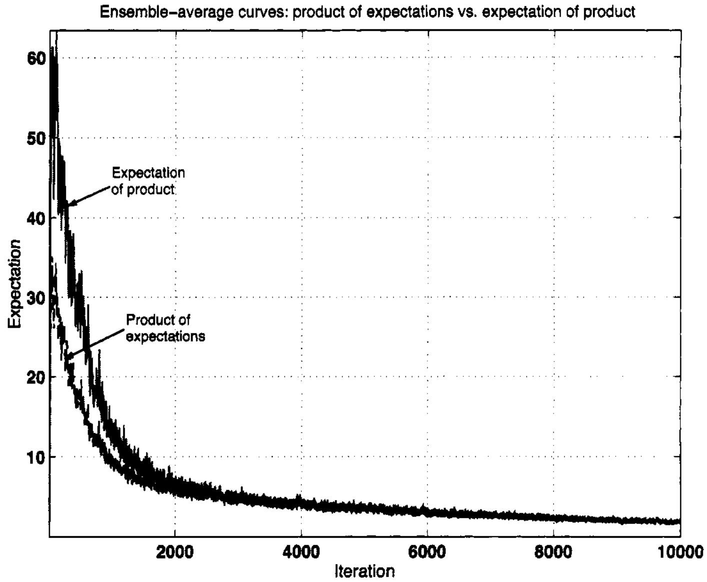

FIGURE 16.1 Ensemble-average curves for the expectation of the product, $\mathsf{E}\left(\|u_i\|^2|e_a(i)|^2\right)$ (*upper curve*), and for the product of expectations, $\left(\mathsf{E}\,\|u_i\|^2\right)\cdot\left(\mathsf{E}\,|e_a(i)|^2\right)$ (*lower curve*), for a 20-tap LMS filter with step-size $\mu = 0.001$. The curves are obtained by averaging over 1000 experiments.

It is seen that both curves tend to each other as time progresses so that it is reasonable to use the separation assumption (16.7) to approximate the expectation of the product as the product of expectations.

Now substituting (16.9) into (16.5) leads to the following expression for the EMSE of LMS:

$$\boxed{\zeta^{\mathrm{LMS}} = \frac{\mu\sigma_v^2\mathrm{Tr}(R_u)}{2-\mu\mathrm{Tr}(R_u)}} \qquad \text{(over a wider range of } \mu) \tag{16.10}$$

This result will be revisited in Chapter 23; see the discussion following the statement of Thm. 23.3.

16.4 WHITE GAUSSIAN INPUT

One particular case for which the term $\mathsf{E}\,\|u_i\|^2|e_a(i)|^2$ that appears in (16.5) can be evaluated in closed-form occurs when u_i has a circular Gaussian distribution with a diagonal covariance matrix, say,

$$R_u = \sigma_u^2\mathrm{I}, \quad \sigma_u^2 > 0 \tag{16.11}$$

That is, when the probability density function of u_i is of the form (cf. Lemma A.1):

$$f_u(u) = \frac{1}{\pi^M}\frac{1}{\det R_u}\,\exp\{-uR_u^{-1}u^*\} = \frac{1}{(\pi\sigma_u^2)^M}\,\exp\{-\|u\|^2/\sigma_u^2\}$$

The diagonal structure of R_u amounts to saying that the entries of u_i are uncorrelated among themselves and that each has variance σ_u^2. The analysis can still be carried out in closed form even without this whiteness assumption; it suffices to require the regressors to be Gaussian. Moreover, R_u does not need to be a scaled multiple of the identity. We treat this more general situation in Sec. 23.1; see Prob. IV.19 for further motivation and the discussion following Thm. 23.3.

In addition to (16.11), we shall assume in this subsection that

$$\boxed{\text{At steady state, } \tilde{\boldsymbol{w}}_{i-1} \text{ is independent of } \boldsymbol{u}_i} \tag{16.12}$$

Conditions (16.11) and (16.12) enable us to evaluate $\mathsf{E}\,\|\boldsymbol{u}_i\|^2|\boldsymbol{e}_a(i)|^2$ explicitly. Before doing so, however, it is worth pointing out that to perform this task, it has been common in the literature to rely not on (16.12) but instead on a set of conditions known collectively as the *independence assumptions*. These assumptions require the data $\{\boldsymbol{d}(i), \boldsymbol{u}_i\}$ to satisfy the following conditions:

(i) The sequence $\{\boldsymbol{d}(i)\}$ is i.i.d.

(ii) The sequence $\{\boldsymbol{u}_i\}$ is also i.i.d.

(iii) Each $\boldsymbol{u}_i$ is independent of previous $\{\boldsymbol{d}_j, j < i\}$.

(iv) Each $\boldsymbol{d}(i)$ is independent of previous $\{\boldsymbol{u}_j, j < i\}$.

(v) The $\boldsymbol{d}(i)$ and $\boldsymbol{u}_i$ are jointly Gaussian.

(vi) In the case of complex-valued data, the $\boldsymbol{d}(i)$ and $\boldsymbol{u}_i$ are individually and jointly circular random variables, i.e., they satisfy $\mathsf{E}\,\boldsymbol{u}_i^\mathsf{T}\boldsymbol{u}_i = 0$, $\mathsf{E}\,\boldsymbol{d}^2(i) = 0$, and $\mathsf{E}\,\boldsymbol{u}_i^\mathsf{T}\boldsymbol{d}(i) = 0$.

The independence assumptions (i)–(vi) are in general restrictive since, in practice, the sequence $\{\boldsymbol{u}_i\}$ is rarely i.i.d. Consider, for example, the case in which the regressors $\{\boldsymbol{u}_i\}$ correspond to state vectors of an FIR implementation, as in the channel estimation application of Sec. 10.5. In this case, two successive regressors share common entries and cannot be statistically independent. Still, when the step-size is sufficiently small, the conclusions that are obtained under the independence assumptions (i)–(vi) tend to be realistic — see App. 24.A. This may explain their widespread use in the adaptive filtering literature. While restrictive, they provide significant simplifications in the derivations and tend to lead to results that match reasonably well with practice for small step-sizes. The key question of course is how *large* the step-size can be in order to validate the conclusions of an independence analysis. There does not seem to exist a clear answer to this inquiry in the literature.

Condition (16.12) is less restrictive than the independence assumptions (i)–(vi). Actually, assumption (16.12) is implied by the independence conditions. To see this, recall from the discussion in Sec. 15.2 that $\tilde{\boldsymbol{w}}_{i-1}$ is a function of the variables $\{\boldsymbol{w}_{-1}; \boldsymbol{d}(i-1), \ldots, \boldsymbol{d}(0); \boldsymbol{u}_{i-1}, \ldots, \boldsymbol{u}_0\}$. Therefore, if the sequence $\boldsymbol{u}_i$ is assumed i.i.d., and if $\boldsymbol{u}_i$ is independent of all previous $\{\boldsymbol{d}(j)\}$ and of $\boldsymbol{w}_{-1}$, then $\boldsymbol{u}_i$ will be independent of $\tilde{\boldsymbol{w}}_{i-1}$ *for all* i. Note, in addition, that condition (16.12) is only requiring the independence of $\{\tilde{\boldsymbol{w}}_{i-1}, \boldsymbol{u}_i\}$ to hold in *steady-state*; which is a considerably weaker assumption than what is implied by the full blown independence assumptions (i)–(vi). Moreover, assumption (16.12) is reasonable for small step-sizes μ. Intuitively, this is because the update term in (15.24) is relatively small for small μ and the statistical dependence of $\tilde{\boldsymbol{w}}_{i-1}$ on $\boldsymbol{u}_i$ becomes weak. Furthermore, in steady-state, the error $\boldsymbol{e}(i)$ is also small, which makes the update term in (15.24) even smaller.

Remark 16.1 (Independence) In our development in this chapter we do *not* adopt the independence assumptions (i)–(vi). Instead, we rely almost exclusively on the separation condition (16.7). It is only in this subsection, for Gaussian regressors, that we also use assumption (16.12) in order to show how to re-derive some known results for Gaussian data from the variance relation (15.40). ◇

So let us return to the term $\mathsf{E}\,\|\boldsymbol{u}_i\|^2|\boldsymbol{e}_a(i)|^2$ in (16.5) and show how it can be evaluated under (16.12), and under the assumption of circular Gaussian regressors. First we show how to express $\mathsf{E}\,\|\boldsymbol{u}_i\|^2|\boldsymbol{e}_a(i)|^2$ in terms of $\mathsf{E}\,|\boldsymbol{e}_a(i)|^2$ (see (16.17) further ahead). Thus, note the following sequence of identities:

$$\begin{aligned}
\mathsf{E}\,\|\boldsymbol{u}_i\|^2|\boldsymbol{e}_a(i)|^2 &= \mathsf{E}\left(\boldsymbol{u}_i\boldsymbol{u}_i^*(\boldsymbol{u}_i\tilde{\boldsymbol{w}}_{i-1}\tilde{\boldsymbol{w}}_{i-1}^*\boldsymbol{u}_i^*)\right)\\
&= \mathsf{E}\ \mathsf{Tr}\left(\boldsymbol{u}_i\boldsymbol{u}_i^*\boldsymbol{u}_i\tilde{\boldsymbol{w}}_{i-1}\tilde{\boldsymbol{w}}_{i-1}^*\boldsymbol{u}_i^*\right)\\
&= \mathsf{E}\ \mathsf{Tr}\left(\boldsymbol{u}_i^*\boldsymbol{u}_i\tilde{\boldsymbol{w}}_{i-1}\tilde{\boldsymbol{w}}_{i-1}^*\boldsymbol{u}_i^*\boldsymbol{u}_i\right)\\
&= \mathsf{Tr}\,\mathsf{E}\left(\boldsymbol{u}_i^*\boldsymbol{u}_i\tilde{\boldsymbol{w}}_{i-1}\tilde{\boldsymbol{w}}_{i-1}^*\boldsymbol{u}_i^*\boldsymbol{u}_i\right)
\end{aligned}\tag{16.13}$$

where in the second equality we used the fact that the trace of a scalar is equal to the scalar itself, and in the third equality we used the property that $\mathsf{Tr}(AB) = \mathsf{Tr}(BA)$ for any matrices A and B of compatible dimensions.

We now evaluate the term $\mathsf{E}\left(\boldsymbol{u}_i^*\boldsymbol{u}_i\tilde{\boldsymbol{w}}_{i-1}\tilde{\boldsymbol{w}}_{i-1}^*\boldsymbol{u}_i^*\boldsymbol{u}_i\right)$, which is a covariance matrix. To do so, we recall the following property of conditional expectations, namely, that for any two random variables $\boldsymbol{x}$ and $\boldsymbol{y}$, it holds that $\mathsf{E}\,\boldsymbol{x} = \mathsf{E}\left(\mathsf{E}\,[\boldsymbol{x}|\boldsymbol{y}]\right)$ — see (1.4). Therefore, in steady-state,

$$\begin{aligned}
\mathsf{E}\left(\boldsymbol{u}_i^*\boldsymbol{u}_i\tilde{\boldsymbol{w}}_{i-1}\tilde{\boldsymbol{w}}_{i-1}^*\boldsymbol{u}_i^*\boldsymbol{u}_i\right) &= \mathsf{E}\left[\mathsf{E}\left(\boldsymbol{u}_i^*\boldsymbol{u}_i\tilde{\boldsymbol{w}}_{i-1}\tilde{\boldsymbol{w}}_{i-1}^*\boldsymbol{u}_i^*\boldsymbol{u}_i\mid\boldsymbol{u}_i\right)\right]\\
&= \mathsf{E}\left[\boldsymbol{u}_i^*\boldsymbol{u}_i\ \mathsf{E}\left(\tilde{\boldsymbol{w}}_{i-1}\tilde{\boldsymbol{w}}_{i-1}^*\mid\boldsymbol{u}_i\right)\boldsymbol{u}_i^*\boldsymbol{u}_i\right]\\
&= \mathsf{E}\left(\boldsymbol{u}_i^*\boldsymbol{u}_iC_{i-1}\boldsymbol{u}_i^*\boldsymbol{u}_i\right)
\end{aligned}\tag{16.14}$$

where in the last step we used assumption (16.12), namely, that $\tilde{\boldsymbol{w}}_{i-1}$ and $\boldsymbol{u}_i$ are independent so that

$$\mathsf{E}\left(\tilde{\boldsymbol{w}}_{i-1}\tilde{\boldsymbol{w}}_{i-1}^*\mid\boldsymbol{u}_i\right) = \mathsf{E}\,\tilde{\boldsymbol{w}}_{i-1}\tilde{\boldsymbol{w}}_{i-1}^* \overset{\Delta}{=} C_{i-1}$$

We are also denoting the covariance matrix of $\tilde{\boldsymbol{w}}_{i-1}$ by C_{i-1}. We do not need to know the value of C_{i-1}, as the argument will demonstrate — see the remark following (16.17). We are then reduced to evaluating the expression $\mathsf{E}\,\boldsymbol{u}_i^*\boldsymbol{u}_iC_{i-1}\boldsymbol{u}_i^*\boldsymbol{u}_i$. Due to the circular Gaussian assumption on $\boldsymbol{u}_i$, this term has the same form as the general term that we evaluated earlier in Lemma A.3 for Gaussian variables, with the identifications

$$\boldsymbol{z} \leftarrow \boldsymbol{u}_i^*, \quad W \leftarrow C_{i-1}, \quad \Lambda \leftarrow \sigma_u^2\mathrm{I}$$

so that we can use the result of that lemma to write

$$\mathsf{E}\,\boldsymbol{u}_i^*\boldsymbol{u}_iC_{i-1}\boldsymbol{u}_i^*\boldsymbol{u}_i = \sigma_u^4\left[\mathsf{Tr}(C_{i-1})\mathrm{I} + C_{i-1}\right]\tag{16.15}$$

Substituting this equality into (16.13) we obtain

$$\mathsf{E}\,\|\boldsymbol{u}_i\|^2|\boldsymbol{e}_a(i)|^2 = \mathsf{Tr}\left[\sigma_u^4\mathsf{Tr}(C_{i-1})\mathrm{I} + \sigma_u^4C_{i-1}\right] = (M+1)\sigma_u^4\mathsf{Tr}(C_{i-1})\tag{16.16}$$

Now repeating the same argument that led to (16.14), we also find that

$$\begin{aligned}
\mathsf{E}\,|\boldsymbol{e}_a(i)|^2 = \mathsf{E}\left(\boldsymbol{u}_i\tilde{\boldsymbol{w}}_{i-1}\tilde{\boldsymbol{w}}_{i-1}^*\boldsymbol{u}_i^*\right) &= \mathsf{E}\left(\boldsymbol{u}_iC_{i-1}\boldsymbol{u}_i^*\right)\\
&= \mathsf{E}\left(\boldsymbol{u}_iC_{i-1}\boldsymbol{u}_i^*\right)\\
&= \mathsf{Tr}\,\mathsf{E}\left(\boldsymbol{u}_i^*\boldsymbol{u}_iC_{i-1}\right)\\
&= \mathsf{Tr}\,\mathsf{E}\left(R_uC_{i-1}\right)\\
&= \sigma_u^2\mathsf{Tr}\,\mathsf{E}\left(C_{i-1}\right)
\end{aligned}$$

This expression relates $\text{Tr}(C_{i-1})$ to $\mathsf{E}\,|e_a(i)|^2$. Substituting into (16.16) we obtain

$$\mathsf{E}\,\|\boldsymbol{u}_i\|^2|e_a(i)|^2 \;=\; (M+1)\sigma_u^2\mathsf{E}\,|e_a(i)|^2 \tag{16.17}$$

This relation expresses the desired term $\mathsf{E}\,\|\boldsymbol{u}_i\|^2|e_a(i)|^2$ as a scaled multiple of $\mathsf{E}\,|e_a(i)|^2$ alone — observe that C_{i-1} is cancelled out. Using this result in (16.5), we get

$$\boxed{\zeta^{\text{LMS}} = \frac{\mu M\sigma_v^2\sigma_u^2}{2-\mu(M+1)\sigma_u^2} \quad \text{(for complex-valued data)}} \tag{16.18}$$

The above derivation assumes complex-valued data. If the data were real-valued, then the same arguments would still apply with the only exception of Lemma A.3. Instead, we would employ the result of Lemma A.2 and replace (16.15) by

$$\mathsf{E}\,\boldsymbol{u}_i^{\mathsf{T}}\boldsymbol{u}_iC_{i-1}\boldsymbol{u}_i^{\mathsf{T}}\boldsymbol{u}_i \;=\; \sigma_u^4\left[\text{Tr}(C_{i-1})\mathrm{I}+2C_{i-1}\right]$$

with an additional scaling factor of 2 (now $C_{i-1} = \mathsf{E}\,\tilde{\boldsymbol{w}}_{i-1}\tilde{\boldsymbol{w}}_{i-1}^{\mathsf{T}}$). Then (16.16) and (16.17) would become

$$\mathsf{E}\,\|\boldsymbol{u}_i\|^2e_a^2(i) = (M+2)\sigma_u^4\text{Tr}(C_{i-1}) = (M+2)\sigma_u^2\mathsf{E}\,e_a^2(i) \tag{16.19}$$

and the resulting expression for the EMSE is

$$\boxed{\zeta^{\text{LMS}} = \frac{\mu M\sigma_v^2\sigma_u^2}{2-\mu(M+2)\sigma_u^2} \quad \text{(for real-valued data)}} \tag{16.20}$$

16.5 STATEMENT OF RESULTS

We summarize the earlier results for the LMS filter in the following statement. A conclusion that stands out from the expressions in the lemma is that the performance of LMS is dependent on the filter length M, the step-size μ, and the input covariance matrix R_u.

Lemma 16.1 (EMSE of LMS) Consider the LMS recursion (16.1) and assume the data $\{\boldsymbol{d}(i), \boldsymbol{u}_i\}$ satisfy model (15.16). Then its EMSE can be approximated by the following expressions:

1. For sufficiently small step-sizes, it holds that $\zeta^{\text{LMS}} = \mu\sigma_v^2\text{Tr}(R_u)/2$.
2. Under the separation assumption (16.7), it holds that

$$\zeta^{\text{LMS}} = \mu\sigma_v^2\text{Tr}(R_u)/[2-\mu\text{Tr}(R_u)]$$

3. If $\boldsymbol{u}_i$ is Gaussian with $R_u = \sigma_u^2\mathrm{I}$, and under the steady-state assumption (16.12), it holds that

$$\zeta^{\text{LMS}} = \mu M\sigma_v^2\sigma_u^2/[2-\mu(M+\gamma)\sigma_u^2]$$

where $\gamma = 2$ if the data is real-valued and $\gamma = 1$ if the data is complex-valued and $\boldsymbol{u}_i$ circular. Here M is the dimension of $\boldsymbol{u}_i$.

In all cases, the misadjustment is obtained by dividing the EMSE by σ_v^2.

16.6 SIMULATION RESULTS

Figures 16.2–16.4 show the values of the steady-state MSE of a 10-tap LMS filter for different choices of the step-size and for different signal conditions. The theoretical values are obtained by using the expressions from Lemma 16.1. For each step-size, the experimental value is obtained by running LMS for 4×10^5 iterations and averaging the squared-error curve $\{|e(i)|^2\}$ over 100 experiments in order to generate the ensemble-average curve. The average of the last 5000 entries of the ensemble-average curve is then used as the experimental value for the MSE. The data $\{\boldsymbol{d}(i), \boldsymbol{u}_i\}$ are generated according to model (15.16) using Gaussian noise with variance $\sigma_v^2 = 0.001$.

In Fig. 16.2, the regressors $\{\boldsymbol{u}_i\}$ do not have shift structure (i.e., they do not correspond to regressors that arise from a tapped-delay-line implementation). The regressors are generated as independent realizations of a Gaussian distribution with a covariance matrix R_u whose eigenvalue spread is $\rho = 5$. Observe from the leftmost plot how expression (16.6) leads to a good fit between theory and practice for small step-sizes. On the other hand, as can be seen from the rightmost plot, expression (16.10) provides a better fit over a wider range of step-sizes.

In Fig. 16.3, the regressors $\{\boldsymbol{u}_i\}$ have shift structure and they are generated by feeding correlated data $\{\boldsymbol{u}(i)\}$ into a tapped delay line. The correlated data are obtained by filtering a unit-variance i.i.d. Gaussian random process $\{\boldsymbol{s}(i)\}$ through a first-order auto-regressive model with transfer function

$$\sqrt{1-a^2}/(1-az^{-1})$$

and $a = 0.8$. It is shown in Prob. IV.1 that the auto-correlation sequence of the resulting process $\{\boldsymbol{u}(i)\}$ is

$$r(k) = \mathsf{E}\,\boldsymbol{u}(i)\boldsymbol{u}(i-k) = a^{|k|}$$

for all integer values k. In this way, the covariance matrix R_u of the regressor $\boldsymbol{u}_i$ is a 10×10 Toeplitz matrix with entries $\{a^{|i-j|}, 0 \le i, j \le M-1\}$.

In Fig. 16.4, regressors with shift structure are again used but they are now generated by feeding into the tapped delay line a unit-variance *white* (as opposed to correlated) process so that $R_u = \sigma_u^2 \mathrm{I}$ with $\sigma_u^2 = 1$. This situation allows us to verify the third result in Lemma 16.1. It is seen from all these simulations that the expressions of Lemma 16.1 provide reasonable approximations for the EMSE of the LMS filter. In particular, expression

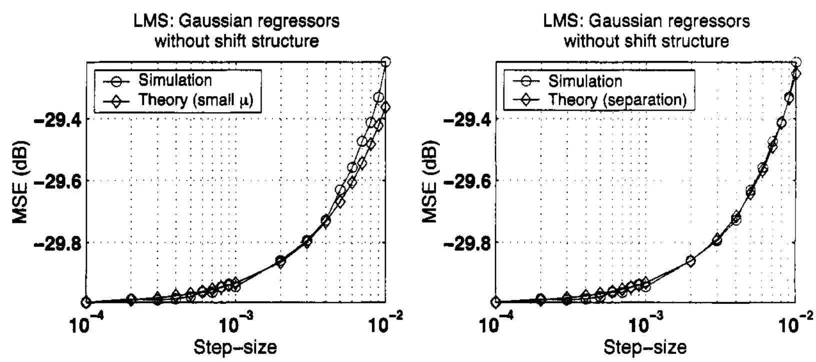

FIGURE 16.2 Theoretical and simulated MSE for a 10-tap LMS filter with $\sigma_v^2 = 0.001$ and Gaussian regressors *without* shift structure. The leftmost plot compares the simulated MSE with expression (16.6), which was derived under the assumption of small step-sizes. The rightmost plot uses expression (16.10), which was derived using the separation assumption (16.7).

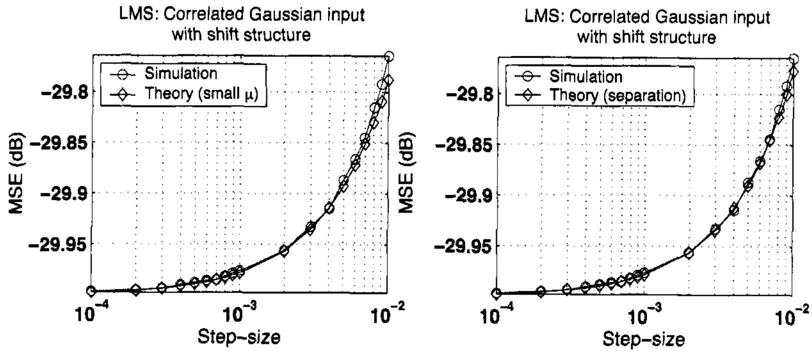

FIGURE 16.3 Theoretical and simulated MSE for a 10-tap LMS filter with $\sigma_v^2 = 0.001$ and regressors *with* shift structure. The regressors are generated by feeding *correlated* data into a tapped delay line. The leftmost plot compares the simulated MSE with expression (16.6), which was derived under the assumption of small step-sizes. The rightmost plot uses expression (16.10), which was derived using the separation assumption (16.7).

(16.10), which was derived under the separation assumption (16.7), provides a good match between theory and practice.

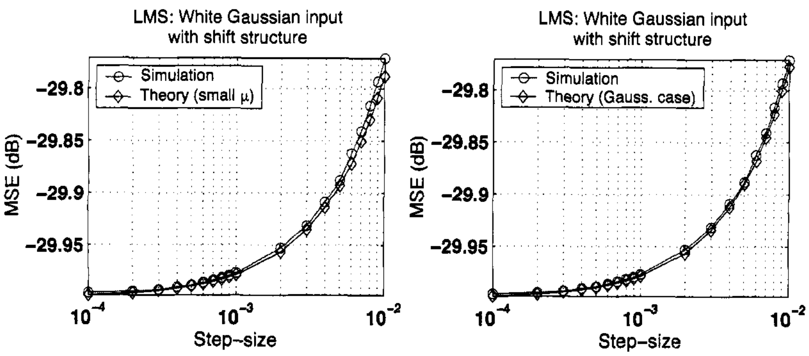

FIGURE 16.4 Theoretical (using (16.20)) and simulated MSE for a 10-tap LMS filter with $\sigma_v^2 = 0.001$ and regressors with shift structure. The regressors are generated by feeding *white* Gaussian input data with unit variance into a tapped delay line.