SIAM J. CONTROL AND OPTIMIZATION
Vol. 33, No. 6, pp. 1926–1951, November 1995

016

THE CONTINUUM-ARMED BANDIT PROBLEM*

RAJEEV AGRAWAL†

Abstract. In this paper we consider the multiarmed bandit problem where the arms are chosen from a subset of the real line and the mean rewards are assumed to be a continuous function of the arms. The problem with an infinite number of arms is much more difficult than the usual one with a finite number of arms because the built-in learning task is now infinite dimensional. We devise a kernel estimator-based learning scheme for the mean reward as a function of the arms. Using this learning scheme, we construct a class of certainty equivalence control with forcing schemes and derive asymptotic upper bounds on their learning loss. To the best of our knowledge, these bounds are the strongest rates yet available. Moreover, they are stronger than the $o(n)$ required for optimality with respect to the average-cost-per-unit-time criterion.

Key words. bandit problems, controlled i.i.d. process, stochastic adaptive control, certainty equivalence with forcing, learning loss, continuous arms

AMS subject classifications. Primary, 93E35, 62G20, 62L05; Secondary, 60F15, 60F25.

1. Introduction.

In this paper we consider the multiarmed bandit problem where the arms are chosen from a subset of the real line and the mean rewards are assumed to be a continuous function of the arms. Prior work on the multiarmed bandit problem has dealt almost exclusively with only a finite number of independent arms. One of the early papers on this topic was by Robbins [26] who constructs a *consistent* (or optimal average-reward-per-unit-time) policy. More recently, the seminal work of Lai and Robbins [23], [22] addressed this problem with the stronger *learning loss* criterion (defined in (5)). They obtained asymptotic lower bounds on the learning loss and constructed *asymptotically efficient* schemes that achieved those bounds. Various extensions of the basic Lai and Robbins formulation have been obtained by Anantharam, Varaiya, and Walrand [6], [7]; by Agrawal, Hegde, and Teneketzis [2], [3]; and by Agrawal, Teneketzis, and Anantharam [5], [4]. In [5] the arms are allowed to be dependent, and the dependence is explicitly exploited to improve the performance. One of the few papers that does deal with an infinite set of arms is by Yakowitz and Lowe [28]. They consider only the *ϵ-learning loss* criterion (defined in (6)), and for this criterion they get a weaker rate than the one obtained in this paper (Corollary 5.4). However, unlike us, they do not make any continuity assumptions.

Note that the case of an infinite number of arms is much more difficult than the usual one of a finite number of arms because the built-in learning task is now infinite dimensional whereas previously it was only finite dimensional. In this paper we exploit the continuity of the mean reward as a function of the arms to devise a class of learning schemes based on kernel estimators. We obtain an upper bound on the almost sure and L^p uniform consistency rates for these estimators. These bounds strengthen the ones available in the nonparametric regression literature as detailed in §3, and may thus be of independent interest.

Subsequently, using the approach taken in [4], we construct a class of adaptive control schemes based on certainty equivalence control with forcing and derive asymptotic upper bounds on their learning loss. These bounds are not only much stronger

* Received by the editors September 16, 1992; accepted for publication (in revised form) July 11, 1994. This research was supported by National Science Foundation grant ECS-8919818.

† Department of Electrical and Computer Engineering, University of Wisconsin–Madison, Madison, Wisconsin 53706-1691 (agrawal@engr.wisc.edu).

than the $o(n)$ required for optimality with respect to the average-cost-per-unit-time criterion, but are also, to the best of our knowledge, the best rates available to date.

The rest of the paper is organized as follows: In §2 we give the precise problem formulation. In §3 we concentrate solely on the learning aspect of the problem. We construct a class of learning schemes and derive an upper bound on its rate of convergence (Corollary 3.4). In §4 we derive various limit laws for "moving averages" that are needed to obtain the rates of convergence in §3. Finally, in §5 we construct a class of adaptive control schemes based on the learning schemes of §3, and obtain upper bounds on their learning loss (Corollaries 5.2 and 5.4).

2. The problem. Consider a (memoryless) discrete-time stochastic system modeled by a controlled i.i.d. process, i.e.,

$$\begin{aligned} P(X_n \in B | U_1, X_1, \ldots, U_{n-1}, X_{n-1}, U_n = u) &= P(X_n \in B | U_n = u) \\ &= P(X_1 \in B | U_1 = u) \end{aligned} \tag{1}$$

where $\{U_n, X_n\}_{n=1}^{\infty}$ is the chronological sequence of controls and states. The states X_n take values in some arbitrary set $\mathcal{X}$, and the controls U_n are chosen from a bounded set $\mathcal{U} \subset \mathbb{R}$. In particular, we will assume that $\mathcal{U} = [\Delta, 1-\Delta]$ for some $0 < \Delta < 1/2$; any arbitrary bounded subset of $\mathbb{R}$ can be handled easily by a slight modification. There is a one-step reward, $r(U_n, X_n)$, associated with each pair $(U_n, X_n), n \geq 1$, where $r : \mathcal{U} \times \mathcal{X} \to \mathbb{R}$. Let $m : \mathcal{U} \to \mathbb{R}$ be defined by

$$m(u) := E[r(U_n, X_n) | U_n = u] = E[r(u, X_1) | U_1 = u],$$

and let $W_n := r(U_n, X_n) - m(U_n)$. Then we can write

$$r(U_n, X_n) = m(U_n) + W_n,$$

where $E[W_n | U_n = u] = 0$ and

$$\begin{aligned} P(W_n \in B | U_1, W_1, \ldots, U_{n-1}, W_{n-1}, U_n = u) &= P(W_n \in B | U_n = u) \\ &= P(W_1 \in B | U_1 = u). \end{aligned} \tag{2}$$

Throughout the rest of the paper we shall assume that $\{U_n, W_n\}$ satisfy the following condition. There exist $\varsigma, s_0 > 0$ such that

$$E[\exp(sW_n) | U_n = u] \leq \exp(\varsigma^2 s^2/2) \quad \forall |s| \leq s_0, u \in \mathcal{U}. \tag{3}$$

In that case define

$$\sigma := \inf\{\varsigma > 0 : \text{ there exists } s_0 > 0 \text{ such that (3) holds}\}. \tag{4}$$

Then $E[|W_n|^2 | U_n = u] \leq \sigma^2$ for all $u \in \mathcal{U}$.

The problem is to design an adaptive control scheme $\gamma = \{\gamma_n\}_{n=1}^{\infty}$, i.e., $U_n = \gamma_n(U_1, X_1, \ldots, U_{n-1}, X_{n-1})$, so as to "maximize" the total reward

$$J_n^{\gamma} := \sum_{i=1}^{n} r(U_i, X_i) = \sum_{i=1}^{n} m(U_i) + \sum_{i=1}^{n} W_i\ ,$$

as $n \to \infty$. First note that

$$J_n^{\gamma} \leq nm^* + \sum_{i=1}^{n} W_i,$$

where $m^* := \sup_{u \in \mathcal{U}} m(u)$. Now if $m : \mathcal{U} \to \mathbb{R}$ were known, then for any constant $M > 0$, we could construct a scheme γ^M such that

$$J_n^{\gamma^M} \geq nm^* + \sum_{i=1}^{n} W_i - M.$$

In the absence of knowledge of m it is desirable to approach this performance as closely as possible. For this purpose define the *learning loss*

$$L_n := \sum_{i=1}^{n} m^* - m(U_i) \ . \tag{5}$$

Also define the ϵ-*learning loss*

$$L_n^{\epsilon} := \sum_{i=1}^{n} I\{m^* - m(U_i) > \epsilon\} \ . \tag{6}$$

Therefore, more precisely, the problem is to design adaptive control schemes for which the learning loss (or ϵ-learning loss) increases slowly regardless of the actual m. In this paper we will investigate almost sure and L^p $(p \geq 1)$ rates for the learning loss and L^1 rate for the ϵ-learning loss.

Throughout the rest of the paper we shall assume that $m : \mathcal{U} \to \mathbb{R}$ is uniformly locally Lipschitz with constant L $(0 \leq L < \infty)$, exponent α $(0 < \alpha \leq 1)$, and restriction δ $(\delta > 0)$, i.e.,

$$u, u' \in \mathcal{U}, |u' - u| \leq \delta \quad \Rightarrow \quad |m(u') - m(u)| \leq L|u' - u|^{\alpha}.$$

Let $ulL(\alpha, L, \delta)$ denote the class of all such functions.

3. The learning scheme. In this section we concentrate solely on the learning aspect of the problem as a first step toward the construction of adaptive control schemes. More precisely, we are interested in choosing the controls $\{U_n\}$ and in constructing the estimates $\{U_n^*\}$ based on the observations $\{m(U_n) + W_n\}$ made at those points, so that $m^* - m(U_n^*)$ converges to 0 as rapidly as possible. Note that the estimates U_n^* need not be the same as the control values U_n. In this section we construct a class of learning schemes and obtain bounds on their rates of convergence. These bounds are important because they determine precisely how the learning schemes can be used to design good adaptive control schemes.

First note that the only assumption we have made on the function m is that it is uniformly locally Lipschitz. Since we do not make any unimodality assumptions we will need a "global search" strategy. Moreover, since the function m may not be differentiable, we cannot use algorithms that rely on the estimation of the gradient. For both these reasons, the Kiefer–Wolfowitz-type (K–W-type) stochastic approximation algorithm is not appropriate for the problem at hand. Even with the unimodality assumption, we would require additional differentiability conditions (existence, boundedness, continuity of the second derivative) in order to get any rate of convergence results for the K–W-type algorithms (see Fabian [12], Nevel'son and Has'minskii [25], Kushner and Clark [21]). Furthermore, the rates of convergence obtained for these algorithms are slower than the ones obtained for the algorithm used in this paper. This is possibly due to the fact that we are interested in constructing the estimates U_n^* so as to minimize the difference from the maximum, $m^* - m(U_n^*)$, rather than to minimize the actual distance $|u^* - U_n^*|$ from the point of maximum u^*. In fact, a point of maximum need not even exist for the problem considered in this paper.

Recently, several researchers have studied a variant of the K–W-type gradient algorithm, which incorporates the global search aspects of simulated annealing algorithms by adding a slowly decreasing noise term. For this algorithm, convergence to the set of global maxima can be established without any unimodality assumption, but with suitable differentiability conditions (see [15], [20], [8], [14]). However, to the best of our knowledge, no results are available on the rate of convergence of these algorithms.

The learning schemes constructed in this section are based on the approach used by Devroye [11]. We first obtain a "uniformly good" estimate $\hat{m}_n$ of the function m, and then use the point of maximum of $\hat{m}_n$ as the estimate U_n^*. We use a nearly equispaced control (*design*) sequence $\{u_n\}$ defined below in (7)–(8). This is a natural choice that corresponds to a progressively finer sampling of $[0,1]$ along the dyadic rationals. Note that the equispaced design scheme itself is not a *progressive* design scheme, i.e., the set of design points at stage n is not a subset of the design points at stage $n+1$. The nearly equispaced design scheme described below is the best progressive approximation to the equispaced design scheme in the sense that it visits the equispaced design at stages $n = 2^m, m = 0, 1, \ldots,$ which occur earlier than the successive visits of any other progressive design scheme.[1] Based on the observations obtained at these points, we estimate the function m by means of a kernel estimator defined below in (9)–(11). In Theorem 3.1 we obtain an upper bound (with an explicit constant) on the almost sure and L^p $(p \geq 1)$ uniform consistency rates of this estimation scheme. In Theorem 3.2 we show that the above upper bound is also a lower bound on the in-probability uniform consistency rates of this estimation scheme for the i.i.d. noise case. This shows us that the rate and associated constant obtained in Theorem 3.1 cannot be improved in general.

The problem of estimating the function $m : \mathcal{U} \to \mathbb{R}$ on the basis of "noisy" measurements $\{m(U_n)+W_n\}_{n=1}^{\infty}$ taken at a sequence of points $\{U_n\}_{n=1}^{\infty} \in \mathcal{U}$ has been extensively investigated in the nonparametric regression literature in statistics. While the almost sure rate itself is well known in the nonparametric regression literature, the associated constant is not (see Stone [27]; Mack and Silverman [24]; Härdle and Luckhaus [18]; Härdle, Janssen and Serfling [17]). Also see Härdle [16, Chap. 4, pp. 89–98] for excellent coverage of results available to date on this nonparametric regression problem. The identification of a sharp constant associated with this rate is an important and challenging problem, as evidenced by the work of Fabian [13], who obtains a constant associated with the in-probability rate for a nonprogressive design scheme with a piecewise polynomial estimator. In fact, Fabian comments that specifying the constants in the order of convergence seems difficult for the estimates considered up to now. One of the key contributions of this section is to provide precisely such a constant for a kernel estimator, which is one of the popular estimators considered in the nonparametric regression literature. As explained in §4, we crucially exploit the structure of the progressive nearly equispaced scheme to obtain the sharp constant associated with the almost sure rate. The constant corresponding to the design/estimation scheme considered in this paper is better than the constant obtained

[1] There are other progressive design schemes which visit the equispaced design scheme at stages $n = 2^m, m = 0, 1, \ldots.$ The design scheme described in this paper is the "best" progressive approximation to the equispaced design scheme in possibly a much stronger sense.

by Fabian [13] for his design/estimation scheme (see the second remark following Corollary 3.3). Another contribution of the results obtained in this section is to establish that the same rates also hold both almost surely and in L^p, whereas all of the papers cited above obtain the rates only in probability or almost surely.

Note that the uniform error between the estimate and the true function m can easily be decomposed into the sum of a bias term and a variance term (see (14) below). The bias term is controlled by the continuity of the function m, while the variance term is controlled by the "noise" in the measurements and by the choice of the design scheme. The variance term is essentially a "moving average." Our main result in Theorem 3.1 on the rate of convergence makes use of some fundamental limit theorems on these moving averages that are developed in §4.

We use the nearly equispaced design sequence $\{u_n\}_{n=1}^{\infty} \in \mathcal{U} \subsetneq [0,1]$ described below. For $n = 1, 2, \ldots$, consider the binary representation of $n-1$:

$$n - 1 = \ldots b_3 b_2 b_1. \tag{7}$$

First, choose the points $\tilde{u}_n \in [0,1]$ to be dyadic rationals with the binary representation:

$$\tilde{u}_n = 0.b_1 b_2 b_3 \ldots. \tag{8}$$

Note that this is the best progressive approximation to the equispaced design scheme on $[0,1]$. The actual design points, $u_n \in \mathcal{U} = [\Delta, 1-\Delta] \subset [0,1]$, are obtained by projecting $\tilde{u}_n$ onto the control set $\mathcal{U}$. Thus

$$u_n := (\tilde{u}_n \wedge (1-\Delta)) \vee \Delta.$$

The reason for choosing the actual design points in this manner is to ensure that asymptotically we get enough observations close to the two boundaries. If we had chosen $\mathcal{U} = [0,1]$ and $u_n = \tilde{u}_n$, then we would have gotten only half as many observations in a window centered at one of the boundary points as we would in the interior. Note that under any deterministic design scheme (such as the one above), $\{W_n\}$ are independent by (2).

We use a window estimator which is a special case of the more general class of Nadaraya–Watson kernel estimators to estimate m. Thus,

$$\hat{m}_n(u) = \frac{\sum_{i=1}^{n} K_{h_n}(u - \tilde{u}_i)(m(u_i) + W_i)}{\sum_{i=1}^{n} K_{h_n}(u - \tilde{u}_i)}, \tag{9}$$

where

$$K_{h_n}(u) = h_n^{-1} K(u/h_n), \tag{10}$$

$$K(u) = I\{|u| \le 1/2\} \tag{11}$$

is the window kernel, and $\{h_n\}$ is a sequence of bandwidths to be specified later. Finally, we choose the estimate U_n^* by

$$U_n^* := \operatorname*{argmax}_{u \in \mathcal{U}} \hat{m}_n(u).$$

Note that $\hat{m}_n$ is a piecewise constant function with at most a finite number of discontinuities. Hence, there exists at least one such argmax. Ties may be resolved in some fixed but arbitrary measurable manner. This completes the description of our class of learning schemes.

In the remaining part of this section we obtain bounds on the rate of convergence of the above scheme: in Theorem 3.1 and Corollary 3.3 for the estimates $\hat{m}_n$, and in Corollary 3.4 for U_n^*. Let

$$d_\infty(\hat{m}, m) = \sup_{u \in \mathcal{U}} |\hat{m}(u) - m(u)|$$

be the uniform metric. The following theorem obtains (1) almost sure and L^p rates of convergence for $\sup_{m \in ulL(\alpha,L,\delta)} d_\infty(\hat{m}_n, m)$, along with a sharp constant, and (2) a large deviations-type result for $P(\sup_{m \in ulL(\alpha,L,\delta)} d_\infty(\hat{m}_n, m) > \epsilon)$ for any $\epsilon > 0$.

THEOREM 3.1. *For the above window estimator with any bandwidth sequence, h_n, such that* (i) *h_n is nonincreasing,* (ii) *nh_n is nondecreasing, and* (iii) *$A'n^{a'} \leq nh_n \leq An^a$ for some $0 < a' \leq a < 1$, $A', A > 0$, we have*

$$\overline{\lim_{n\to\infty}} \frac{\sup_{m \in ulL(\alpha,L,\delta)} d_\infty(\hat{m}_n, m)}{\frac{L}{\alpha+1}\left(\frac{h_n}{2}\right)^\alpha + \left(\frac{2\log(1/h_n)}{nh_n}\right)^{1/2}\sigma} \leq 1 \quad a.s. \text{ and in } L^p, p \geq 1. \tag{12}$$

Also, for the above window estimator with any bandwidth sequence, h_n, such that (i) *$h_n \to 0$ as $n \to \infty$, and* (ii) *$nh_n \geq A'n^{a'}$ for some $0 < a' < 1$, $A' > 0$, we have*

$$\overline{\lim_{\epsilon \searrow 0}} \frac{1}{\epsilon^2} \overline{\lim_{n\to\infty}} \frac{1}{nh_n} \log P\left(\sup_{m \in ulL(\alpha,L,\delta)} d_\infty(\hat{m}_n, m) > \epsilon\right) \leq -\frac{1}{8\sigma^2}. \tag{13}$$

Remark. The above theorem can easily be adapted to handle $\mathcal{U} = [a,b]$ for any $-\infty < a < b < \infty$. For each $m : [a,b] \to \mathbb{R}$, $m \in ulL(\alpha, L, \delta)$, simply consider the shifted and rescaled function $\tilde{m} : [\Delta, 1-\Delta] \to \mathbb{R}$ defined by $\tilde{m}(u) = m(a + (u - \Delta)(b-a)(1-2\Delta)^{-1})$. Note that

$$m \in ulL(\alpha, L, \delta) \Leftrightarrow \tilde{m} \in ulL\left(\alpha, L\left(\frac{b-a}{1-2\Delta}\right)^\alpha, \delta\left(\frac{1-2\Delta}{b-a}\right)\right).$$

Next observe that (12) holds for any $\delta > 0$ and $0 < \Delta < 1/2$; however, it does explicitly depend on α and L. Hence, in light of the above discussion, it would be desirable to choose Δ arbitrarily small.

Proof. Since $nh_n \geq A'n^{a'}$ for some $0 < a' < 1$, $A' > 0$, we can pick n_0 such that for all $n \geq n_0$, $\sum_{i=1}^n I\{|u - \tilde{u}_i| \leq h_n/2\} > 0 \quad \forall u \in \mathcal{U}$. Then, for all $n \geq n_0$,

$$\begin{aligned}\hat{m}_n(u) &= \frac{\sum_{i=1}^n K_{h_n}(u - \tilde{u}_i)(m(u_i) + W_i)}{\sum_{i=1}^n K_{h_n}(u - \tilde{u}_i)} \\ &= \frac{\sum_{i=1}^n h_n^{-1} I\{|u - \tilde{u}_i| \leq h_n/2\}(m(u_i) + W_i)}{\sum_{i=1}^n h_n^{-1} I\{|u - \tilde{u}_i| \leq h_n/2\}} \\ &= \frac{\sum_{i=1}^n I\{|u - \tilde{u}_i| \leq h_n/2\}(m(u_i) + W_i)}{\sum_{i=1}^n I\{|u - \tilde{u}_i| \leq h_n/2\}}.\end{aligned}$$

Thus,

$$\begin{aligned}|\hat{m}_n(u) - m(u)| \leq &\frac{\sum_{i=1}^n I\{|u - \tilde{u}_i| \leq h_n/2\}|m(u_i) - m(u)|}{\sum_{i=1}^n I\{|u - \tilde{u}_i| \leq h_n/2\}} \\ &+ \frac{|\sum_{i=1}^n I\{|u - \tilde{u}_i| \leq h_n/2\} W_i|}{\sum_{i=1}^n I\{|u - \tilde{u}_i| \leq h_n/2\}}.\end{aligned} \tag{14}$$

Since $h_n \to 0$ as $n \to \infty$, we can pick $n_1 \geq n_0$ so that $h_n/2 \leq \min\{\delta, \Delta\}$ for all $n \geq n_1$. Then, for any $u \in \mathcal{U} = [\Delta, 1-\Delta]$ and $n \geq n_1$, we have

$$b_n := nh_n - \lfloor \log_2 n \rfloor - 1 \leq \sum_{i=1}^{n} I\{|u - \tilde{u}_i| \leq h_n/2\} \leq nh_n + \lfloor \log_2 n \rfloor + 1 =: a_n. \tag{15}$$

Thus, by the condition $nh_n \geq A'n^{a'}$ for some $0 < a' < 1$, $A' > 0$, it follows that $\sum_{i=1}^{n} I\{|u - \tilde{u}_i| \leq h_n/2\} \approx a_n \approx b_n \approx nh_n$, where by $x_n \approx y_n$ we mean that $x_n/y_n \to 1$ as $n \to \infty$.

Also, for any $m \in ulL(\alpha, L, \delta)$ and $n \geq n_1$, we have

$$\begin{aligned} &\sum_{i=1}^{n} I\{|u - \tilde{u}_i| \leq h_n/2\}|m(u_i) - m(u)| \\ &\qquad \leq \sum_{i=1}^{n} I\{|u - \tilde{u}_i| \leq h_n/2\} L|\tilde{u}_i - u|^{\alpha} \\ &\qquad \leq \frac{L}{\alpha+1}\left(\frac{h_n}{2}\right)^{\alpha} [nh_n + 2(\alpha+1)(\lfloor \log_2 n \rfloor + 1)], \end{aligned} \tag{16}$$

where the second inequality is established in the appendix. Combining this with (15), we get for any $m \in ulL(\alpha, L, \delta)$ and $n \geq n_1$,

$$\begin{aligned} \frac{\sum_{i=1}^{n} I\{|u - \tilde{u}_i| \leq h_n/2\}|m(u_i) - m(u)|}{\sum_{i=1}^{n} I\{|u - \tilde{u}_i| \leq h_n/2\}} &\leq \frac{L}{\alpha+1}\left(\frac{h_n}{2}\right)^{\alpha} \frac{1 + 2(\alpha+1)\frac{\lfloor \log_2 n \rfloor + 1}{nh_n}}{1 - \frac{\lfloor \log_2 n \rfloor + 1}{nh_n}} \\ &= \frac{L}{\alpha+1}\left(\frac{h_n}{2}\right)^{\alpha} (1 + \gamma_n) \quad \text{(say)} \\ &\approx \frac{L}{\alpha+1}\left(\frac{h_n}{2}\right)^{\alpha}. \end{aligned} \tag{17}$$

Let $\{\tilde{W}_i^n\}_{i=1}^n$ be a permutation of $\{W_i\}_{i=1}^n$ arranged in order of $\{\tilde{u}_i\}_{i=1}^n$. Define the partial sums of $\{\tilde{W}_i^n\}_{i=1}^n$ as follows:

$$\tilde{S}_{j,k}^n := \sum_{i=j+1}^{k} \tilde{W}_i^n. \tag{18}$$

Also let

$$\tilde{B}_{n,a_n} := \max_{\substack{0 \leq j \leq k \leq n \\ k-j \leq a_n}} |\tilde{S}_{j,k}^n|. \tag{19}$$

Then

$$\left| \sum_{i=1}^{n} I\{|u - \tilde{u}_i| \leq h_n/2\} W_i \right| \leq \tilde{B}_{n,a_n}. \tag{20}$$

Putting together (14)–(20) we get that for all $n \geq n_1$,

$$\begin{aligned} \sup_{m \in ulL(\alpha,L,\delta)} d_{\infty}(\hat{m}_n, m) &= \sup_{m \in ulL(\alpha,L,\delta)} \sup_{u \in \mathcal{U}} |\hat{m}_n(u) - m(u)| \\ &\leq \frac{L}{\alpha+1}\left(\frac{h_n}{2}\right)^{\alpha} (1 + \gamma_n) + \frac{\tilde{B}_{n,a_n}}{b_n}. \end{aligned} \tag{21}$$

Asymptotic bounds on the "moving average" $\tilde{B}_{n,a_n}$ are obtained in §4. In particular, by Theorems 4.5 and 4.7 we know that $\overline{\lim}_n \tilde{B}_{n,a_n}/\beta_n \le \sigma$ a.s. and in L^p, $p \ge 1$, respectively, where $\beta_n = (2a_n(\log(n/a_n) + \log\log n))^{1/2}$, provided a_n satisfies some conditions. It is straightforward to check that $a_n = nh_n + \lfloor \log_2 n \rfloor + 1$ satisfies the conditions of Theorems 4.5 and 4.7 whenever h_n satisfies the conditions of this theorem. Finally, we have

$$\begin{aligned}\beta_n/b_n &= \frac{(2a_n(\log(n/a_n) + \log\log n))^{1/2}}{nh_n - \lfloor \log_2 n \rfloor - 1} \\ &\approx \frac{(2nh_n(\log(n/nh_n) + \log\log n))^{1/2}}{nh_n} \\ &\approx \left(\frac{2\log(1/h_n)}{nh_n}\right)^{1/2}. \end{aligned} \tag{22}$$

Combining all of the above, we get (12).

We will now prove the remaining part of the theorem. Let $0 < \eta < \epsilon$ be fixed but arbitrary. Since the bandwidth sequence $h_n \to 0$ and $\gamma_n \to 0$ as $n \to 0$, we can pick $n_2 \ge n_1$ so that $h_n/2 \le \min\{\delta, \Delta, (\eta(\alpha+1)/2L(1+\gamma_n))^{1/\alpha}\}$, for all $n \ge n_2$. Also, by the condition that $nh_n \ge A'n^{a'}$ for some $0 < a' < 1$, $A' > 0$, it follows that $b_n/a_n \to 1$ as $n \to \infty$. Thus, we can pick $n_3 \ge n_2$ so that we have $(\epsilon - \eta/2)b_n \ge (\epsilon - \eta)a_n$ for all $n \ge n_3$. Then, from (21), we have for all $n \ge n_3$

$$\begin{aligned} P\left(\sup_{m \in ulL(\alpha,L,\delta)} d_\infty(\hat{m}_n, m) > \epsilon\right) &\le P\left(\frac{L}{\alpha+1}(h_n/2)^\alpha + \tilde{B}_{n,a_n}/b_n > \epsilon\right) \\ &\le P(\tilde{B}_{n,a_n}/b_n > \epsilon - \eta/2) \\ &\le P(\tilde{B}_{n,a_n}/a_n > \epsilon - \eta). \end{aligned} \tag{23}$$

Substituting n, m, p, t, s, a in Lemma 4.1 by $n, a_n, \lfloor nd/a_n \rfloor + 1, (\epsilon - 2\eta), \eta, (\epsilon - 2\eta - 2\rho)/(\epsilon - 2\eta)$, respectively, we have

$$\begin{aligned} P(\tilde{B}_{n,a_n}/a_n > \epsilon - \eta) \le \frac{1}{1-c}\left(\frac{nd}{a_n} + 1\right) \Bigg[&\max_{0 \le j \le n} P(|\tilde{S}^n_{j,j+a_n}|/a_n > (\epsilon - 2\eta - 2\rho)) \\ &+ \max_{0 \le j \le n} P(|\tilde{S}^n_{j,j+\lfloor a_n/d \rfloor + 1}|/a_n > \rho) \Bigg] \end{aligned} \tag{24}$$

where c is such that

$$\max_{0 \le j \le n} \max_{1 \le k \le a_n+1} P(|\tilde{S}^n_{j,j+k}|/a_n > \eta/4) \le c < 1. \tag{25}$$

By Chebyshev's inequality we have

$$\max_{0 \le j \le n} \max_{1 \le k \le a_n+1} P(|\tilde{S}^n_{j,j+k}|/a_n > \eta/4) \le \frac{\sigma^2(a_n+1)}{(a_n\eta/4)^2}.$$

By the assumptions on h_n, $a_n = nh_n + \lfloor \log_2 n \rfloor + 1 \to \infty$ as $n \to \infty$. Therefore, given any $0 < c < 1$, there exists $n_4 \ge n_3$ such that (25) holds for all $n \ge n_4$. We now apply Lemma 4.2 to upper bound the right-hand side (RHS) of (24). Given any $\varsigma > \sigma$, let s_0 be such that (30) holds. Let $\epsilon \le \varsigma^2 s_0$. Then by Lemma 4.2 we have

$$P(|\tilde{S}^n_{j,j+a_n}|/a_n > (\epsilon - 2\eta - 2\rho)) \le 2\exp\left(-\frac{(\epsilon - 2\eta - 2\rho)^2 a_n^2}{2a_n\varsigma^2}\right).$$

Similarly, if $d\rho \le \varsigma^2 s_0$, then by Lemma 4.2 we have

$$P(|\tilde{S}^n_{j,j+\lfloor a_n/d\rfloor+1}|/a_n > \rho) \le 2\exp\left(-\frac{d\rho^2 a_n^2}{2a_n\varsigma^2}\right).$$

Thus, by choosing $d = 4$ and $\rho = \epsilon/4$, we get that for all $\epsilon \le \varsigma^2 s_0$, $\eta < \epsilon/4$, and $n \ge n_4$,

$$P(\tilde{B}_{n,a_n}/a_n > \epsilon - \eta) \le \frac{1}{1-c}\left(\frac{nd}{a_n}+1\right)4\exp\left(-\frac{((\epsilon/2)-2\eta)^2 a_n}{2\varsigma^2}\right).$$

Thus, by (23) and the above, it follows that

$$\begin{aligned}\overline{\lim_{n\to\infty}}\frac{1}{a_n}\log P\left(\sup_{m\in ulL(\alpha,L,\delta)} d_\infty(\hat{m}_n,m) > \epsilon\right) &\le \overline{\lim_{n\to\infty}}\frac{\log(1/h_n)}{nh_n+\lfloor\log_2 n\rfloor+1} - \frac{((\epsilon/2)-2\eta)^2}{2\varsigma^2}\\ &= -\frac{((\epsilon/2)-2\eta)^2}{2\varsigma^2}.\end{aligned}$$

Note that the last equality follows from the assumptions on h_n. The left-hand side (LHS) above does not depend on η. By letting $\eta \to 0$, we get for all $\epsilon \le \varsigma^2 s_0$,

$$\overline{\lim_{n\to\infty}}\frac{1}{a_n}\log P\left(\sup_{m\in ulL(\alpha,L,\delta)} d_\infty(\hat{m}_n,m) > \epsilon\right) \le -\frac{\epsilon^2}{8\varsigma^2}.$$

Now dividing by ϵ^2, taking limits as $\epsilon \to 0$, and finally letting $\varsigma \to \sigma$, we obtain (13). □

Below, we show that the rate and the associated constant identified in (12) in the previous theorem are the best possible ones.

THEOREM 3.2. *If, in addition to the conditions already imposed on* $\{W_n\}$, *we assume that*

$$P(W_n \in B|U_n = u') = P(W_1 \in B|U_1 = u), \quad \forall B,\ u', u,\ n,$$

then we can also show that (12) *holds with equality in probability, i.e.,*

$$\lim_{n\to\infty} P\left(\frac{\sup_{m\in ulL(\alpha,L,\delta)} d_\infty(\hat{m}_n,m)}{\frac{L}{\alpha+1}(\frac{h_n}{2})^\alpha + (\frac{2\log(1/h_n)}{nh_n})^{1/2}\sigma} < 1-\epsilon\right) = 0 \quad \forall\epsilon > 0.$$

Proof. This follows because by choosing $m \in ulL(\alpha,L,\delta)$ to be given by $m(u) = \pm|u-c|^\alpha$ for $c \in [\Delta', 1-\Delta']$ with $\Delta < \Delta' < 1/2$, one can show that there exists an n_0' such that for all $n \ge n_0'$,

$$\begin{aligned}\sup_{m\in ulL(\alpha,L,\delta)} d_\infty(\hat{m}_n,m) \ge \sup_{c\in[\Delta',1-\Delta']}\Bigg[&\frac{\sum_{i=1}^n I\{|c-\tilde{u}_i| \le h_n/2\}L|\tilde{u}_i - c|^\alpha}{\sum_{i=1}^n I\{|c-\tilde{u}_i|\le h_n/2\}}\\ &+ \frac{|\sum_{i=1}^n I\{|c-\tilde{u}_i|\le h_n/2\}W_i|}{\sum_{i=1}^n I\{|c-\tilde{u}_i|\le h_n/2\}}\Bigg].\end{aligned} \tag{26}$$

Corresponding to the upper bound on the first term on the RHS used earlier in (16), we have the following lower bound which is also established in the appendix. For all $n \ge 1$ and $u \in \mathcal{U}$,

$$\sum_{i=1}^n I\{|u-\tilde{u}_i| \le h_n/2\}L|\tilde{u}_i - u|^\alpha \ge \frac{L}{\alpha+1}\left(\frac{h_n}{2}\right)^\alpha [nh_n - 2(\alpha+1)(\lfloor\log_2 n\rfloor+1)]. \tag{27}$$

Using (27) and (15) in (26) we get for all $n \geq n_0'$,

$$\begin{aligned}\sup_{m\in ulL(\alpha,L,\delta)} d_\infty(\hat{m}_n, m) \geq \frac{L}{\alpha+1}\left(\frac{h_n}{2}\right)^\alpha \frac{1-2(\alpha+1)\frac{\lfloor \log_2 n\rfloor+1}{nh_n}}{1+\frac{\lfloor \log_2 n\rfloor+1}{nh_n}} \\ + \max_{\Delta' n\leq j\leq (1-\Delta')n} \frac{|\tilde{S}^n_{j,j+a_{n,j}}|}{\beta_n}\frac{\beta_n}{a_n} \\ =: r_n + s_n t_n,\end{aligned}$$

where $b_n \leq a_{n,j} \leq a_n$ for all $1 \leq j \leq n$, $n = 1, 2, \ldots$. As shown earlier in (17) and (22), for any $\epsilon > 0$ there exists an $n_1' \geq n_0'$ such that for all $n \geq n_1'$,

$$r_n \geq (1-\epsilon)\frac{L}{\alpha+1}\left(\frac{h_n}{2}\right)^\alpha$$

and

$$t_n \geq (1-\epsilon)^{1/2}\left(\frac{2\log(1/h_n)}{nh_n}\right)^{1/2}.$$

Thus, for all $n \geq n_1'$ we get

$$\begin{aligned}&P\left(\frac{\sup_{m\in ulL(\alpha,L,\delta)} d_\infty(\hat{m}_n, m)}{\frac{L}{\alpha+1}(\frac{h_n}{2})^\alpha + (\frac{2\log(1/h_n)}{nh_n})^{1/2}\sigma} < 1-\epsilon\right) \\ &\leq P\left(r_n + s_n t_n < (1-\epsilon)\left(\frac{L}{\alpha+1}\left(\frac{h_n}{2}\right)^\alpha + \left(\frac{2\log(1/h_n)}{nh_n}\right)^{1/2}\sigma\right)\right) \\ &\leq P(s_n < (1-\epsilon)^{1/2}\sigma)\end{aligned}$$

which goes to 0 as $n \to \infty$ by Theorem 4.3. Note that the collection of random variables $\{V_i^n\}$ obtained by setting $\{V_i^n\}_{i=1}^n = \{\tilde{W}_i^n\}_{i=1}^n$ for $n = 1, 2, \ldots$ satisfies conditions (C1)–(C4), but not (C5) (see §4). However, since the RHS above does not depend on (C5), we can still apply Theorem 4.3, whereas we would need condition (C5) to obtain the corresponding almost sure result from Theorem 4.3. □

COROLLARY 3.3. *For the above window estimator with the bandwidth sequence* $h_n = h(\log n/n)^{1/(2\alpha+1)}$, $h > 0$, *we have*

$$\varlimsup_{n\to\infty} \frac{\sup_{m\in ulL(\alpha,L,\delta)} d_\infty(\hat{m}_n, m)}{(\frac{\log n}{n})^{\alpha/(2\alpha+1)}} \leq c(\alpha, L, \sigma, h) \quad \text{a.s. and in } L^p, p \geq 1, \tag{28}$$

where

$$c(\alpha, L, \sigma, h) := \frac{L}{\alpha+1}\left(\frac{h}{2}\right)^\alpha + \left(\frac{2}{h(2\alpha+1)}\right)^{1/2}\sigma. \tag{29}$$

Proof. First note that the above choice of h_n satisfies the conditions of Theorem 3.1. Moreover,

$$\frac{L}{\alpha+1}\quad\left(\frac{h_n}{2}\right)^\alpha + \left(\frac{2\log(1/h_n)}{nh_n}\right)^{1/2}\sigma$$

$$= \frac{L}{\alpha+1}\left(\frac{h}{2}\right)^{\alpha}\left(\frac{\log n}{n}\right)^{\alpha/(2\alpha+1)}$$
$$+\left[\frac{-2\log h + \frac{2}{(2\alpha+1)}(\log n - \log\log n)}{hn^{2\alpha/(2\alpha+1)}(\log n)^{1/(2\alpha+1)}}\right]^{1/2}\sigma$$
$$\approx \left[\frac{L}{\alpha+1}\left(\frac{h}{2}\right)^{\alpha} + \sigma\left(\frac{2}{h(2\alpha+1)}\right)^{1/2}\right]\left(\frac{\log n}{n}\right)^{\alpha/(2\alpha+1)}. \qquad \square$$

Remark. Note that the bandwidth sequence $\{h_n\}$ chosen above minimizes the rate of convergence given by (12). This is so because if $h_n/(\log n/n)^{1/(2\alpha+1)} \to 0$ or ∞, then

$$\left[\frac{L}{\alpha+1}\left(\frac{h_n}{2}\right)^{\alpha} + \left(\frac{2\log(1/h_n)}{nh_n}\right)^{1/2}\sigma\right]\left(\frac{\log n}{n}\right)^{\alpha/(2\alpha+1)} \to \infty.$$

Moreover, the constant $c(\alpha, L, \sigma, h)$ associated with the above rate is minimized by choosing

$$h^* = 2(\sigma(\alpha+1)/2\alpha L(2\alpha+1)^{1/2})^{2/(2\alpha+1)}.$$

The minimum value of the above constant, say $c^*(\alpha, L, \sigma)$, is given by

$$c^*(\alpha, L, \sigma) = \left[\frac{L}{\alpha+1}\sigma^{2\alpha}(2\alpha+1)^{-\alpha}\right]^{1/(2\alpha+1)}[(2\alpha)^{-2\alpha/(2\alpha+1)} + (2\alpha)^{1/(2\alpha+1)}].$$

Remark. The above constant can be compared with the one obtained by Fabian [13]. Note that while we consider uniformly locally Lipschitz functions with exponent $0 < \alpha \leq 1$, [13] considers functions with bounded rth derivative, $r \geq 1$. Thus $\alpha = r = 1$ is the only case in common between the two papers. In particular, for $\alpha = 1$ and $\sigma = 1$, our constant reduces to

$$c^*(1, L, 1) = L^{1/3}\frac{(1/4)^{1/3} + 2^{1/3}}{6^{1/3}}.$$

In comparison, the optimal constant in [13, Thm. 5.4, p. 1359] for the above case ($r = 1, s = 2$) is greater by a factor of $\pi^{1/3}$. Also note that while our results hold both almost surely and in $L^p, p \geq 1$, those of [13] hold only in probability.

The following corollary establishes the same rates for the sequence U_n^*.

COROLLARY 3.4. *For the learning scheme constructed above with the bandwidth sequence $h_n = h(\log n/n)^{1/(2\alpha+1)}$, $h > 0$, we have*

$$\overline{\lim_{n\to\infty}} \frac{\sup_{m\in ulL(\alpha,L,\delta)}(m^* - m(U_n^*))}{(\frac{\log n}{n})^{\alpha/(2\alpha+1)}} \leq 2c(\alpha, L, \sigma, h) \quad \text{a.s. and in } L^p, p \geq 1,$$

where $c(\alpha, L, \sigma, h)$ is given by (29). *Also, for the the same learning scheme with any bandwidth sequence, h_n, such that* (i) *$h_n \to 0$ as $n \to \infty$, and* (ii) *$nh_n \geq A'n^{a'}$ for some $0 < a' < 1$, $A' > 0$, we have*

$$\overline{\lim_{n\to\infty}} \frac{1}{nh_n} \log P\left(\sup_{m\in ulL(\alpha,L,\delta)}(m^* - m(U_n^*)) > \epsilon\right) \leq -\Gamma(\epsilon/2)$$

for all $\epsilon > 0$, for some $\Gamma(\epsilon) > 0$.

Proof. Both parts of the corollary follow immediately from Theorem 3.1, Corollary 3.3, and the following observation. Since, $\hat{m}_n(U_n^*) = \max_{u\in\mathcal{U}} \hat{m}_n(u)$,

$$m(u) - m(U_n^*) \le \hat{m}_n(u) + d_\infty(\hat{m}_n, m) - \hat{m}_n(U_n^*) + d_\infty(\hat{m}_n, m) \le 2d_\infty(\hat{m}_n, m).$$

Taking $\sup_{u\in\mathcal{U}}$ on both sides we get $m^* - m(U_n^*) \le 2d_\infty(\hat{m}_n, m)$. □

4. Limit laws for moving averages. In this section we obtain almost sure and L^p limit theorems for the moving averages that comprise the variance term in the analysis of the kernel estimator considered in §3. If the noise were i.i.d. and the design scheme were equispaced (see conditions (C4) and (C5) below), then exact almost sure limit theorems are available in the literature for these moving averages (see Theorem 4.3). These results can be viewed as generalizations of the law of iterated logarithm. In light of the paragraph following Theorem 4.3, the weaker in-probability result also holds for the the i.i.d. noise case with the progressive nearly equispaced design scheme employed in this paper. This identifies the desired rate and constant. However, in order to get the stronger almost sure results for the progressive nearly equispaced design scheme, we have to work considerably harder. We first obtain a coarser constant in Theorem 4.4 and then recover the desired constant in Theorem 4.5, by applying that result on an appropriately chosen subsequence. This subsequence argument relies heavily on the exact structure of the nearly equispaced design scheme. Thus, the specific design scheme plays a crucial role in the analysis, if we desire to get a sharp constant.

It is also worth noting that the techniques used here to obtain the abovementioned limit theorems parallel those of de Acosta and Kuelbs [10], by relying on classical probability inequalities rather than on the strong approximations approach of Csörgő and Révész [9]. This allows us to obtain results for $\{W_i\}$ which are independent but not necessarily identically distributed. Also, it gives us a way of extending these results to higher dimensions (obtained in a forthcoming paper [1]), for which the strong approximation results are not yet available, to the best of our knowledge (see Mack and Silverman [24]).

In Theorem 4.7 we establish the same rates (with the same constant) in L^p. This is done by means of a device due to Hoffman-Jørgensen [19]. The proof of Theorem 4.7 can be used more generally for establishing L^p counterparts of in-probability rates.

Throughout this section we will consider a collection of random variables $\{V_i^n : n = 1, 2, \ldots; i = 1, 2, \ldots, n\}$ satisfying some of the following conditions.

(C1) $E[V_i^n] = 0$ for all $n = 1, 2, \ldots; i = 1, 2, \ldots, n$.

(C2) $\{V_i^n\}_{i=1}^n$ is an independent collection for each $n = 1, 2, \ldots$.

(C3) There exist $\varsigma, s_0 > 0$ such that

$$E[\exp(sV_i^n)] \le \exp(\varsigma^2 s^2/2) \quad \forall |s| \le s_0,\ n = 1, 2, \ldots; i = 1, 2, \ldots, n. \tag{30}$$

In that case define

$$\sigma := \inf\{\varsigma > 0 : \text{ there exists } s_0 > 0 \text{ such that (30) holds}\}. \tag{31}$$

(C4) $\{V_i^n : n = 1, 2, \ldots; i = 1, 2, \ldots, n\}$ are identically distributed.

(C5) $V_i^n = V_i^i$ for all $n = 1, 2, \ldots; i = 1, 2, \ldots, n$.

Note that under (C1) and (C3), it follows that $E[|V_i^n|^2] \le \sigma^2$, and under (C1), (C3), and (C4), $E[|V_i^n|^2] = \sigma^2$. Also, note that the collection of random variables $\{\tilde{W}_i^n : n = 1, 2, \ldots; i = 1, 2, \ldots, n\}$ obtained in the previous section with $\{\tilde{W}_i^n\}_{i=1}^n$

being a permutation of $\{W_i\}_{i=1}^n$ arranged in order of $\{\tilde{u}_i\}_{i=1}^n$, $n = 1, 2, \ldots$, satisfies conditions (C1), (C2), and (C3).

For each $n = 1, 2, \ldots$, define the partial sums of the random variables $\{V_i^n\}_{i=1}^n$ by

$$S_{j,k}^n := \sum_{i=j+1}^{k} V_i^n. \tag{32}$$

For $m \leq n$, define

$$B_{n,m} := \max_{0 \leq j \leq n} \max_{1 \leq k \leq (m \wedge n-j)} |S_{j,j+k}^n|. \tag{33}$$

Note that by setting $V_i^n = \tilde{W}_i^n$ of the previous section, we get $S_{j,k}^n = \tilde{S}_{j,k}^n$ and $B_{n,m} = \tilde{B}_{n,m}$ defined in (18) and (19), respectively.

Let $\{a_n\}$ be a sequence of positive integers satisfying some of the following conditions:

(A1) $1 \leq a_n \leq n$,
(A2) a_n is nondecreasing,
(A3) n/a_n is nondecreasing,
(A4) $a_n/\log n \to \infty$ as $n \to \infty$.
(A5) For some $a > 0$, $n^{a-1}a_n = o(n^{\varepsilon})$
(A6) For some $a', A' > 0$, $a_n \geq A'n^{a'}$. for every $\varepsilon > 0$.

Finally, let

$$\beta_n = (2a_n(\log(n/a_n) + \log\log n))^{1/2}.$$

In this section we determine the limiting behavior of $\{B_{n,a_n}/\beta_n\}$.

Below we give two lemmas that will be used to establish the limit theorems. The first is a minor modification of a maximal inequality due to de Acosta and Kuelbs [10, Lem. 3.1].

LEMMA 4.1. *Let $\{V_i\}$ be a sequence of independent random variables and let $S_{j,k} := \sum_{i=j+1}^{k} V_i$. Then for every integer $n \geq 0$, $m \geq 0$, $p \geq 0$, $p < n$, $m \leq n$, and $t > 0$, $s > 0$, $0 < a < 1$,*

$$P(B_{n,m} > t+s) \leq \frac{p}{1-c}\left[\max_{0 \leq j \leq n} P(|S_{j,j+m}| > at) + \max_{0 \leq j \leq n} P\left(|S_{j,j+\lfloor \frac{n}{p} \rfloor+1}| > \frac{(1-a)}{2}t\right)\right],$$

provided

$$\max_{0 \leq j \leq n} \max_{1 \leq k \leq m \vee \lfloor \frac{n}{p} \rfloor+1} P(|S_{j,j+k}| \geq s/4) \leq c < 1.$$

LEMMA 4.2. *Let $\{V_i^n\}$ be a collection of random variables satisfying conditions (C2) and (C3) with σ as defined in (31). Given any $\varsigma > \sigma$, let $s_0 > 0$ be such that (30) holds. Then for all $j \geq 0$, $k \geq 1$, and $0 < a \leq k\varsigma^2 s_0$,*

$$P(|S_{j,j+k}^n| \geq a) \leq 2\exp\left(-\frac{a^2}{2k\varsigma^2}\right).$$

Proof. For any $0 < s \leq s_0$,

$$\begin{aligned} P(S_{j,j+k}^n \geq a) &\leq \exp(-sa)E[\exp(sS_{j,j+k}^n)] \\ &= \exp(-sa)E[\exp(sV_{j+1}^n)] \cdots E[\exp(sV_{j+k}^n)] \\ &\leq \exp\left(\frac{k\varsigma^2 s^2}{2} - sa\right). \end{aligned}$$

The same upper bound also holds for $P(-S^n_{j,j+k} \geq a)$. Optimizing the upper bound over $0 \leq s \leq s_0$, we get the desired result. □

The first theorem from [9] gives almost sure rates for the process $\{B_{n,a_n}\}_{n=1}^{\infty}$ under conditions (C1)–(C5).

THEOREM 4.3 (Csörgő and Révész [9, Thm. 3.1.1]). *Let* $\{V_i^n\}$ *be a be a collection of random variables satisfying conditions* (C1)–(C5) *with variance* σ^2. *Let* $\{a_n\}$ *satisfy conditions* (A1)–(A4). *Then*

$$\overline{\lim_{n\to\infty}} \frac{B_{n,a_n}}{\beta_n} = \sigma \ a.s.$$

If in addition $\{a_n\}$ *satisfies* (A5), *then*

$$\lim_{n\to\infty} \frac{\max_{\Delta n \leq j \leq (1-\Delta)n} |S^n_{j,j+a_{n,j}}|}{\beta_n} = \sigma \ a.s.$$

for all $0 \leq \Delta < 1/2$, $0 < \rho \leq 1$, *and* $a_{n,j}$ *satisfying* $\rho a_n \leq a_{n,j} \leq a_n, \forall 1 \leq j \leq n$, $n = 1, 2, \ldots$.

From the above theorem we immediately have the same rate in probability for any collection $\{V_i^n\}$ satisfying conditions (C1)–(C4) (but not necessarily (C5)), i.e.,

$$\lim_n P(B_{n,a_n}/\beta_n > (\sigma + \epsilon)) = 0, \quad \forall \epsilon > 0, \tag{34}$$

and σ is the smallest constant for which (34) holds. In the next theorem we obtain upper bounds on the in-probability rate (with the above constant) and on the almost sure rate (with a larger constant) for $\{V_i^n\}$ that do not necessarily satisfy conditions (C4) or (C5).

THEOREM 4.4. *Let* $\{V_i^n\}$ *be a collection of random variables satisfying conditions* (C1)–(C3) *with* σ *as defined in* (31). *Let* $\{a_n\}$ *satisfy conditions* (A1)–(A4). *Then* (34) *holds. If in addition* $\{a_n\}$ *satisfies* (A5) *for some* $a > 0$, *then*

$$\overline{\lim_{n\to\infty}} \frac{B_{n,a_n}}{\beta_n} \leq \left(1 + \frac{1}{a}\right)^{1/2} \sigma \ a.s.$$

Proof. The theorem will follow from the Borel–Cantelli lemma if we can show that for all $\varsigma > \sigma$ and $\epsilon > 0$,

$$\sum_n P(B_{n,a_n}/\beta_n > (\rho\varsigma + 4\epsilon)) < \infty, \tag{35}$$

where $\rho = (1 + a^{-1})^{1/2}$. Substituting n, m, p, t, s in Lemma 4.1 by $n, a_n, \lfloor nd/a_n \rfloor + 1, (\rho\varsigma + 3\epsilon)\beta_n, \epsilon\beta_n$, respectively, we have

$$\begin{aligned} P(B_{n,a_n}/\beta_n > (\rho\varsigma + 4\epsilon)) \leq 2\left(\frac{nd}{a_n} + 1\right) \max_{0\leq j\leq n} P(|S^n_{j,j+a_n}|/\beta_n > (\rho\varsigma + \epsilon)) \\ +2\left(\frac{nd}{a_n} + 1\right) \max_{0\leq j\leq n} P(|S^n_{j,j+\lfloor a_n/d\rfloor+1}|/\beta_n > \epsilon) \end{aligned} \tag{36}$$

provided

$$\max_{0\leq j\leq n} \max_{1\leq k\leq a_n+1} P(|S^n_{j,j+k}|/\beta_n \geq \epsilon/4) \leq 1/2. \tag{37}$$

By Chebyshev's inequality we have

$$\max_{0\le j\le n}\max_{1\le k\le a_n+1} P(|S^n_{j,j+k}|/\beta_n \ge \epsilon/4) \le \frac{\sigma^2(a_n+1)}{(\beta_n\epsilon/4)^2}.$$

By the definition of β_n and the assumptions on a_n, $a_n/\beta_n^2 \to 0$ as $n\to\infty$. So there exists an n_0 such that (37) holds for all $n \ge n_0$. We now apply Lemma 4.2 to upper bound the RHS of (36). Let s_0 be such that (30) holds. Note that by condition (A4) on $\{a_n\}$ it follows that $\beta_n/a_n \to 0$ as $n\to\infty$. Hence, there exists an $n_1 \ge n_0$ such that for all $n \ge n_1$, $(\rho\varsigma+\epsilon)\beta_n \le a_n\varsigma^2 s_0$, and consequently for all $n\ge n_1$,

$$\begin{aligned}
P(|S^n_{j,j+a_n}|/\beta_n > (\rho\varsigma+\epsilon)) &\le 2\exp\left(-\frac{(\rho\varsigma+\epsilon)^2\beta_n^2}{2\varsigma^2 a_n}\right)\\
&= 2\exp\left(-\frac{(\rho\varsigma+\epsilon)^2 2a_n(\log(n/a_n)+\log\log n)}{2\varsigma^2 a_n}\right)\\
&= 2\exp(-(\rho')^2(\log(n/a_n)+\log\log n))\\
&= 2\left(\frac{a_n}{n\log n}\right)^{(\rho')^2},
\end{aligned}$$

where $\rho' = \rho+\epsilon/\varsigma$. Thus, for all $n\ge n_1$, the first term on the RHS of (36) is bounded above by

$$(38)\; 2\left(\frac{nd}{a_n}+1\right)\max_{0\le j\le n} P(|S^n_{j,j+a_n}|/\beta_n > (\rho\varsigma+\epsilon)) \le 8d\left(\frac{a_n}{n}\right)^{(\rho')^2-1}(\log n)^{-(\rho')^2}.$$

In view of the additional condition on a_n it is easy to check that

$$\left(\frac{a_n}{n}\right)^{(\rho')^2-1}(\log n)^{-(\rho')^2} \le \left(\frac{a_n}{n}\right)^{(\rho')^2-1} = o(1/n^\gamma),$$

where $\gamma = a((\rho+\epsilon/2\varsigma)^2-1) > 1$. By choosing the constant $d > ((\rho\varsigma+\epsilon)/\epsilon)^2$, the same asymptotic upper bound can be obtained for the second term on the RHS of (36). This establishes (35).

Finally, by setting $\rho = 1$ in (36) through (38), it can easily be seen that (34) holds. □

In the next theorem we obtain the same upper bound as in Theorem 4.3 for the sequence $\{\tilde B_{n,a_n}\}$ that was defined in the previous section.

THEOREM 4.5. *Let* $\{\tilde W^n_i\}$ *be the collection of random variables defined in the previous section with* σ *as defined in* (31). *Let* $\{\tilde B_{n,a_n}\}$ *be as defined in* (19). *Let* $\{a_n\}$ *satisfy conditions* (A1)–(A5) *with* $a>0$ *as in* (A5). *Then*

$$\overline{\lim_{n\to\infty}}\frac{\tilde B_{n,a_n}}{\beta_n} \le \sigma \text{ a.s.}$$

Proof. Let $\{\tilde W^{m,n}_i\}_{i=m}^n$ be a permutation of $\{W_i\}_{i=m}^n$ arranged in order of $\{\tilde u_i\}_{i=m}^n$. Define their partial sums

$$(39)\qquad \tilde S^{m,n}_{j,k} := \sum_{i=j+1}^{k}\tilde W^{m,n}_i, \qquad m-1\le j\le k\le n.$$

Define

$$\tilde{B}^m_{n,a_n} := \max_{m-1\le j\le n}\ \max_{1\le k\le (a_n\wedge n-j)} |\tilde{S}^{m,n}_{j,j+k}|. \tag{40}$$

Then, in terms of our prior notation, $\tilde{B}_{n,a_n} = \tilde{B}^1_{n,a_n}$. The proof of this theorem depends crucially on the following simple observation:

$$\tilde{B}_{n,a_n} \le \tilde{B}_{m,a_n} + \tilde{B}^m_{n,2\frac{n-m}{n}a_n+1} \qquad \forall m \le n.$$

This fact is a consequence of the nearly uniformly spaced design sequence $\{u_n\}$. Now let $n_k = \lfloor \theta^k \rfloor$, $k = 1, 2, \dots$ for some $\theta > 1$. Given any n, let k be such that $n_k \le n < n_{k+1}$. Then by substituting $m = n_k$ above and using the fact that a_n is increasing, we get

$$\tilde{B}_{n,a_n} \le \tilde{B}_{n_k,a_{n_{k+1}}} + \tilde{B}^{n_k}_{n,2\frac{n-n_k}{n}a_n+1}.$$

Now, since $\beta_{n_k} \le \beta_n \le \beta_{n_{k+1}}$, we get

$$\frac{\tilde{B}_{n,a_n}}{\beta_n} \le \frac{\beta_{n_{k+1}}}{\beta_{n_k}}\frac{\tilde{B}_{n_k,a_{n_{k+1}}}}{\beta_{n_{k+1}}} + \frac{\tilde{B}^{n_k}_{n,2\frac{n-n_k}{n}a_n+1}}{\beta_n}.$$

Let $\rho = (1 + a^{-1})^{1/2}$. The theorem will follow if for all $\varsigma > \sigma$ and $0 < \eta < 2\varsigma\rho^2$, we can find a $\theta > 1$ such that

$$\overline{\lim_{k\to\infty}}\ \frac{\beta_{n_{k+1}}}{\beta_{n_k}}\frac{\tilde{B}_{n_k,a_{n_{k+1}}}}{\beta_{n_{k+1}}} \le \varsigma + \eta \text{ a.s.}, \tag{41}$$

and

$$\overline{\lim_{n\to\infty}}\ \frac{\tilde{B}^{n_k}_{n,2\frac{n-n_k}{n}a_n+1}}{\beta_n} \le \eta \text{ a.s.} \tag{42}$$

Choose $\theta = 1 + (\eta/\sqrt{2}\rho\varsigma)^2$. It is easy to check that

$$\overline{\lim_{k\to\infty}}\ \frac{\beta_{n_{k+1}}}{\beta_{n_k}} \le \theta.$$

Hence, (41) will follow by Borel–Cantelli if we establish that for all $\epsilon > 0$,

$$\sum_k P\left(\frac{\tilde{B}_{n_{k-1},a_{n_k}}}{\beta_{n_k}} > \varsigma + 4\epsilon\right) < \infty. \tag{43}$$

Using Lemmas 4.1 and 4.2, as was done in the proof of Theorem 4.4 (cf. (36), (38)), we get

$$P\left(\frac{\tilde{B}_{n_{k-1},a_{n_k}}}{\beta_{n_k}} > \varsigma + 4\epsilon\right) \le O((\log n_k)^{-(1+\epsilon/\varsigma)}) \le O(k^{-(1+\epsilon/\varsigma)})$$

and (43) holds. Similarly, (42) will follow by Borel–Cantelli if we establish that for all $\epsilon > 0$,

$$\sum_n P\left(\frac{\tilde{B}^{n_k}_{n,2\frac{n-n_k}{n}a_n+1}}{\beta_n} > \eta + 4\epsilon\right) < \infty.$$

Since $\tilde{B}^{n_k}_{n,2\frac{n-n_k}{n}a_n+1} \leq \tilde{B}^{n_k}_{n,2\frac{\theta-1}{\theta}a_n+1}$, the above will follow from

$$\sum_n P\left(\frac{\tilde{B}^{n_k}_{n,2\frac{\theta-1}{\theta}a_n+1}}{\beta_n} > \eta + 4\epsilon\right) < \infty. \tag{44}$$

Again, using Lemmas 4.1 and 4.2, as was done in the proof of Theorem 4.4 (cf. (36), (38)), we get

$$P\left(\frac{\tilde{B}^{n_k}_{n,2\frac{\theta-1}{\theta}a_n+1}}{\beta_n} > \eta + 4\epsilon\right) \leq O\left(\left(\frac{a_n}{n}\right)^{(1+a^{-1})(1+\epsilon/\eta)-1}\right) \leq o(n^{-\gamma})$$

for some $\gamma > 1$. Hence (44) holds. $\square$

We now proceed to obtain the same upper bound as in Theorem 4.3 but in L^p instead of almost surely. We first establish the following lemma which is needed to obtain the L^p upper bound.

LEMMA 4.6. *Let $\{V_i^n\}$ be a collection of random variables satisfying condition* (C2). *If $\{B_{n,a_n}/\beta_n\}_{n=1}^{\infty}$ is stochastically bounded and* $\sup_n E[\sup_{1\leq i\leq n}(|V_i^n|/\beta_n)^{p'}] < \infty$, *then* $\sup_n E[(B_{n,a_n}/\beta_n)^{p'}] < \infty$ *and consequently, $\{(B_{n,a_n}/\beta_n)^p\}_{n=1}^{\infty}$ is uniformly integrable for all $1 \leq p < p'$.*

Proof. The proof of this lemma is based on a technique borrowed from [19]. We first prove the following claim:

$$P(B_{n,a_n} \geq 2t + s) \leq P(B_{n,a_n} \geq t)^2 + P(N_n \geq s),$$

where $N_n := \sup_{1\leq i\leq n} |V_i^n|$. Let

$$B^{l,m}_{n,a_n} := \max_{\substack{l-1\leq j\leq k\leq m\\ k-j\leq a_n}} |S^n_{j,k}|, \qquad 1 \leq l \leq m \leq n.$$

Thus,

$$B_{n,a_n} = B^{1,n}_{n,a_n} \leq B^{1,m-1}_{n,a_n} + |V^n_m| + B^{m+1,n}_{n,a_n} \quad \text{for any } 1 \leq m \leq n. \tag{45}$$

Let T be the stopping time defined by

$$T := \inf\{1 \leq m \leq n : B^{1,m}_{n,a_n} \geq t\},$$

where $\inf\emptyset = \infty$. Then $B^{1,n}_{n,a_n} \geq 2t + s$ implies that $T \leq n$, and so we have

$$\begin{aligned}
P(B_{n,a_n} \geq 2t + s) &= \sum_{m=1}^{n} P(B_{n,a_n} \geq 2t + s, T = m)\\
&\leq \sum_{m=1}^{n} P(B^{1,m-1}_{n,a_n} + |V^n_m| + B^{m+1,n}_{n,a_n} \geq 2t + s, T = m)\\
&\leq \sum_{m=1}^{n} P(B^{m+1,n}_{n,a_n} \geq t + s - N_n, T = m)\\
&\leq \sum_{m=1}^{n} P(B^{m+1,n}_{n,a_n} \geq t, T = m) + P(N_n \geq s)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{m=1}^{n} P(B_{n,a_n}^{m+1,n} \geq t) P(T = m) + P(N_n \geq s) \\
&\leq P(B_{n,a_n}^{1,n} \geq t) \sum_{m=1}^{n} P(T = m) + P(N_n \geq s) \\
&= P(B_{n,a_n} \geq t)^2 + P(N_n \geq s).
\end{aligned}
$$

The first inequality follows from (45), the second from the fact that $T = m$ implies that $B_{n,a_n}^{1,m-1} < t$, and the third by the definition of N_n. The next equality follows from the fact that $\{B_{n,a_n}^{m+1,n} \geq t\}$ and $\{T = m\}$ are independent events.

Pick A so that $P(B_{n,a_n}/\beta_n \geq A) \leq 1/(2.3^{p'})$. Note that this is possible by the assumption that $\{B_{n,a_n}/\beta_n\}_{n=1}^{\infty}$ is stochastically bounded. Now we have

$$
\begin{aligned}
E[(B_{n,a_n}/\beta_n)^{p'}] &= \int_0^{\infty} p' x^{p'-1} P(B_{n,a_n}/\beta_n \geq x)\, dx \\
&\leq (3A)^{p'} + \int_{3A}^{\infty} p' x^{p'-1} P(B_{n,a_n}/\beta_n \geq x)\, dx \\
&\leq (3A)^{p'} + \int_{3A}^{\infty} p' x^{p'-1} P(N_n/\beta_n \geq x/3)\, dx \\
&\qquad + \int_{3A}^{\infty} p' x^{p'-1} P(B_{n,a_n}/\beta_n \geq x/3)^2\, dx \\
&= (3A)^{p'} + 3^{p'} \int_{A}^{\infty} p' x^{p'-1} P(N_n/\beta_n \geq x)\, dx \\
&\qquad + 3^{p'} \int_{A}^{\infty} p' x^{p'-1} P(B_{n,a_n}/\beta_n \geq x)^2\, dx \\
&\leq (3A)^{p'} + 3^{p'} \int_{A}^{\infty} p' x^{p'-1} P(N_n/\beta_n \geq x)\, dx \\
&\qquad + \frac{1}{2} \int_{A}^{\infty} p' x^{p'-1} P(B_{n,a_n}/\beta_n \geq x)\, dx \\
&\leq (3A)^{p'} + 3^{p'} E[(N_n/\beta_n)^{p'}] + \frac{1}{2} E[(B_{n,a_n}/\beta_n)^{p'}].
\end{aligned}
$$

Thus

$$
E[(B_{n,a_n}/\beta_n)^{p'}] \leq 2(3A)^{p'} + 2.3^{p'} E[(N_n/\beta_n)^{p'}] < \infty.
$$

The uniform integrability of $\{(B_{n,a_n}/\beta_n)^p\}_{n=1}^{\infty}$ for all $1 \leq p < p'$ is an immediate consequence. This completes the proof of Lemma 4.6. □

The L^p upper bound is now established in the following theorem.

THEOREM 4.7. *Let $\{V_i^n\}$ be a collection of random variables satisfying conditions* (C1)–(C3) *with σ as defined in* (31). *Let $\{a_n\}$ satisfy conditions* (A1)–(A4) *and* (A6). *Then*

$$
\overline{\lim_{n\to\infty}}\, E[(B_{n,a_n}/\beta_n)^p]^{1/p} \leq \sigma \quad \forall p \geq 1.
$$

Proof. The theorem will follow from the in-probability bound (34) obtained in Theorem 4.4, if we establish uniform integrability of $\{(B_{n,a_n}/\beta_n)^p\}_{n=1}^{\infty}$. In view of Lemma 4.6, it suffices to establish that $\{B_{n,a_n}/\beta_n\}_{n=1}^{\infty}$ is stochastically bounded and that $\sup_n E[\sup_{1\le i\le n}(|V_i^n|/\beta_n)^{p'}] < \infty$ for some $p' > p$. That $\{B_{n,a_n}/\beta_n\}_{n=1}^{\infty}$ is stochastically bounded also follows from the in-probability bound (34) obtained in Theorem 4.4. Finally, using the fact that $E[|X|] \le E[|X|^r]^{1/r}$ for any $r \ge 1$, we get

$$\begin{aligned}
\sup_n E\left[\sup_{1\le i\le n}(|V_i^n|/\beta_n)^{p'}\right] &= \sup_n (1/\beta_n)^{p'} E\left[\sup_{1\le i\le n}|V_i^n|^{p'}\right] \\
&\le \sup_n (1/\beta_n)^{p'} E\left[\sup_{1\le i\le n}|V_i^n|^{p'r}\right]^{1/r} \\
&\le \sup_n (1/\beta_n)^{p'} E\left[\sum_{i=1}^{n}|V_i^n|^{p'r}\right]^{1/r} \\
&\le \sup_n (1/\beta_n)^{p'} \left(\sum_{i=1}^{n} E[|V_i^n|^{p'r}]\right)^{1/r} \\
&\le \sup_n (n/(\beta_n)^{p'r})^{1/r} \sup_{\substack{1\le i\le n \\ 1\le n<\infty}} E[|V_i^n|^{p'r}]^{1/r} \\
&\le \sup_n (n/((2A')^{1/2}n^{a'/2})^{p'r})^{1/r} \sup_{\substack{1\le i\le n \\ 1\le n<\infty}} E[|V_i^n|^{p'r}]^{1/r} \\
&\le (2A')^{-p'/2}(1 \vee n^{1/r-a'p'/2}) \sup_{\substack{1\le i\le n \\ 1\le n<\infty}} E[|V_i^n|^{p'r}]^{1/r} \\
&= (2A')^{-p'/2} \sup_{\substack{1\le i\le n \\ 1\le n<\infty}} E[|V_i^n|^{p'r}]^{1/r} \\
&< \infty
\end{aligned}$$

if we choose $r = 1 \vee 2/a'p'$. Note that we have used the condition that $a_n \ge A'n^{a'}$ to obtain a corresponding bound on β_n in one of the above steps. Also, we have used condition (C3) on the moment-generating function of $\{V_i^n\}$ to deduce the last inequality. This completes the proof of the theorem. □

5. The adaptive control scheme. In this section we construct a class of certainty equivalence control with forcing-type adaptive control schemes based on the learning schemes constructed in §3. Let $\{\tau_i\}_{i=1}^{\infty}$ be a positive integer-valued sequence to be specified later. Define the related sequence $\{t_i\}_{i=1}^{\infty}$ as follows:

$$t_i := 1 + \sum_{k=1}^{i-1}(\tau_k + 1) = \sum_{k=1}^{i-1}\tau_k + i, \quad i \ge 1.$$

At times t_i, $i \ge 1$ use (force) the ith control u_i from the design sequence of the learning scheme such as the one described in the previous section. Let U_i^* be the estimate based on the corresponding observations at times t_k, $1 \le k \le i$, with the design sequence u_k, $1 \le k \le i$. Use the control U_i^* from time $t_i + 1$ to time $t_{i+1} - 1$, i.e., τ_i times. Thus,

$$U_{t_i} = u_i, \quad U_n = U_i^* \quad \text{for} \quad t_i + 1 \le n \le t_{i+1} - 1, \quad i \ge 1.$$

This completes the description of the adaptive control scheme.

Let

$$\kappa(n) := \min\{i : t_i > n\} - 1 = \max\{i : t_i \le n\} = \max\left\{i : \sum_{k=1}^{i-1}\tau_k + i \le n\right\}.$$

The following theorems and their corollaries provide upper bounds on the learning loss associated with the class of schemes constructed above in terms of $\kappa(n)$.

THEOREM 5.1. *Assume that for a certain learning scheme $m^* - m(U_i^*) = O(r_i)$ a.s. (resp., in L^p with $p \geq 1$) for some known sequence r_i. Then for the certainty equivalence control with forcing-type scheme constructed above with the sequence $\tau_i = \lfloor br_i^{-1} \rfloor$ for some $b > 0$, we have $L_n = O(\kappa(n))$ a.s. (resp., in L^p with $p \geq 1$).*

Proof. Since, $m : \mathcal{U} = [\Delta, 1-\Delta] \to \mathbb{R}$ is uniformly locally Lipschitz, it follows that $K := \sup_{u\in\mathcal{U}} m(u) - \inf_{u\in\mathcal{U}} m(u) < \infty$. Thus,

$$\begin{aligned} L_n &= \sum_{l=1}^{n} m^* - m(U_l) \\ &\leq K\kappa(n) + \sum_{i=1}^{\kappa(n)} \sum_{l=t_i+1}^{t_{i+1}-1} m^* - m(U_l) \\ &= K\kappa(n) + \sum_{i=1}^{\kappa(n)} (m^* - m(U_i^*))\tau_i \\ &\leq K\kappa(n) + \sum_{i=1}^{\kappa(n)} (m^* - m(U_i^*))br_i^{-1}. \end{aligned}$$

Now, for the first part we are given that $m^* - m(U_i^*) = O(r_i)$ a.s. Thus, $\overline{\lim}_i (m^* - m(U_i^*))r_i^{-1} \leq C$ a.s. for some $C \geq 0$ that could depend on ω. That is, for all $\epsilon > 0$, there exists an $i_0 \geq 1$, such that for all $i \geq i_0$, $m^* - m(U_i^*) \leq (C+\epsilon)r_i$ a.s. Then, clearly,

$$\begin{aligned} L_n &\leq K\kappa(n) + \sum_{i=1}^{i_0-1} (m^* - m(U_i^*))br_i^{-1} + \sum_{i=i_0}^{\kappa(n)} (m^* - m(U_i^*))br_i^{-1} \\ &\leq K\kappa(n) + M(\epsilon) + (C+\epsilon)b\kappa(n) \qquad \text{a.s.}, \end{aligned}$$

where $M(\epsilon) = \sum_{i=1}^{i_0-1} (m^* - m(U_i^*))br_i^{-1}$ depends on ϵ but not on n. Dividing by $\kappa(n)$ and taking the limit as $n \to \infty$ we get

$$\overline{\lim_n} \frac{L_n}{\kappa(n)} \leq K + (C+\epsilon)b \qquad \text{a.s.}$$

Now, by letting $\epsilon \to 0$ we get

$$\overline{\lim_n} \frac{L_n}{\kappa(n)} \leq K + Cb \qquad \text{a.s.}$$

For the second part we are given that $m^* - m(U_i^*) = O(r_i)$ in L^p with $p \geq 1$. Thus, $\overline{\lim}_i E[|m^* - m(U_i^*)|^p]r_i^{-p} \leq C^p$ for some $C \geq 0$. That is, for all $\epsilon > 0$ there exists an $i_0 \geq 1$, such that for all $i \geq i_0$, $E[|m^* - m(U_i^*)|^p]r_i^{-p} \leq C^p + \epsilon$. Also by the well-known inequality $|x_1 + \cdots + x_k|^p \leq k^{p-1}(|x_1|^p + \cdots + |x_k|^p)$ we get

$$|L_n|^p \leq 2^{p-1}\left(K^p(\kappa(n))^p + (\kappa(n))^{p-1}\sum_{i=1}^{\kappa(n)} |m^* - m(U_i^*)|^p b^p r_i^{-p}\right).$$

Therefore,

$$E[|L_n|^p] \le 2^{p-1}\left(K^p(\kappa(n))^p + (\kappa(n))^{p-1}\sum_{i=1}^{\kappa(n)} E[|m^* - m(U_i^*)|^p]b^p r_i^{-p}\right)$$
$$\le 2^{p-1}\Bigg(K^p(\kappa(n))^p + (\kappa(n))^{p-1}\sum_{i=1}^{i_0-1} E[|m^* - m(U_i^*)|^p]b^p r_i^{-p}$$
$$+ (\kappa(n))^{p-1}\sum_{i=i_0}^{\kappa(n)} E[|m^* - m(U_i^*)|^p]b^p r_i^{-p}\Bigg)$$
$$\le 2^{p-1}(K^p(\kappa(n))^p + (\kappa(n))^{p-1}M(\epsilon) + (C^p+\epsilon)b^p(\kappa(n))^p),$$

where $M(\epsilon) = \sum_{i=1}^{i_0-1} E[|m^* - m(U_i^*)|^p]b^p r_i^{-p}$ depends on ϵ but not on n. Dividing by $(\kappa(n))^p$ and taking the limit as $n \to \infty$ we get

$$\overline{\lim_n} \frac{E[|L_n|^p]}{(\kappa(n))^p} \le 2^{p-1}(K^p + (C^p+\epsilon)b^p).$$

Now, by letting $\epsilon \to 0$ we get

$$\overline{\lim_n} \frac{E[|L_n|^p]}{(\kappa(n))^p} \le 2^{p-1}(K^p + C^p b^p). \qquad \square$$

Note that we can also obtain a constant (in terms of K, C, b) associated with the rate of increase of the learning loss, $\kappa(n)$. Moreover, $\kappa(n)$ itself depends on b. In fact, if $\tau_i \to \infty$ as $i \to \infty$, then asymptotically $\kappa_b(n) \approx \kappa_1(n/b)$, where the subscript on $\kappa(n)$ denotes the dependence on b. We may therefore want to choose b to minimize the rate along with the constant.

COROLLARY 5.2. *For the certainty equivalence control with forcing-type scheme constructed above, with the learning scheme of* §3 *with the bandwidth sequence* $h_i = h(\log i/i)^{1/(2\alpha+1)}$, *and with the sequence* $\tau_i = \lfloor b(i/\log i)^{\alpha/(2\alpha+1)}\rfloor$ *for some* $b > 0$, *we have* $L_n = O(\tilde{B}^{-1}(n))$ *a.s. and in* $L^p, p \ge 1$, *where* $\tilde{B}^{-1} : [0,\infty) \to [e,\infty)$ *is the inverse of the function* $\tilde{B} : [e,\infty) \to [0,\infty)$ *defined by*

$$\tilde{B}(t) := \int_e^t b\left(\frac{s}{\log s}\right)^{\alpha/(2\alpha+1)} ds.$$

Moreover, $\tilde{B}^{-1}(n) = o(n^{\frac{2\alpha+1}{3\alpha+1}+\eta})$ *for all* $\eta > 0$.

Proof. This corollary will follow immediately from Theorems 5.1 and 3.1, if we can show that $\kappa(n) = O(\tilde{B}^{-1}(n))$, and that $\tilde{B}^{-1}(n) = o(n^{\frac{2\alpha+1}{3\alpha+1}+\eta})$ for all $\eta > 0$, when $\tau_i = \lfloor b(i/\log i)^{\alpha/(2\alpha+1)}\rfloor$. To this end, define the functions $v, B : \mathbb{R}^+ \to \mathbb{R}^+$ based on the sequence $\{\tau_i\}$ as follows:

$$v(s) := \tau_{\lceil s\rceil} + 1,$$
$$B(t) := \int_0^t v(s)\, ds.$$

Then

$$B(i) = \sum_{k=1}^{i} \tau_k + i, \quad i = 0, 1, \ldots.$$

Also, observe that (1) $B(t)$ is continuous in $t > 0$, (2) $B(0) = 0$, (3) $B(t)$ is strictly increasing in $t > 0$ (since $v(s) > 0$), and (4) $B(t) \geq t$ (since $v(s) \geq 1$). Hence, given any n, there exists a unique $0 \leq t \leq n$ such that $B(t) = n$. Denote this solution by $B^{-1}(n)$. Then

$$\begin{aligned}\sum_{k=1}^{\lceil B^{-1}(n)\rceil+1-1} \tau_k + \lceil B^{-1}(n)\rceil + 1 &= B(\lceil B^{-1}(n)\rceil) + 1 \\ &\geq B(B^{-1}(n)) + 1 \\ &= n + 1 > n.\end{aligned}$$

Therefore, by the definition of $\kappa(n)$ it follows that $\kappa(n) \leq \lceil B^{-1}(n)\rceil \leq B^{-1}(n)+1$. It is also easy to see that if we have a function $\tilde{v} : \mathbb{R}^+ \to \mathbb{R}^+$ such that $v(s) \geq \tilde{v}(s),\ s > 0$, then $B(t) \geq \tilde{B}(t),\ t > 0$, and hence $B^{-1}(n) \leq \tilde{B}^{-1}(n),\ n \geq 1$. For the sequence $\tau_i = \lfloor b(i/\log i)^{\alpha/(2\alpha+1)}\rfloor$ under consideration, note that $v(s) = \lfloor b(\lceil s\rceil/\log\lceil s\rceil)^{\alpha/(2\alpha+1)}\rfloor + 1 \geq b(s/\log s)^{\alpha/(2\alpha+1)} I\{s \geq e\} =: \tilde{v}(s)$. Thus $\kappa(n) = O(\tilde{B}^{-1}(n))$. It is easy to verify that $\tilde{B}^{-1}(n) = o(n^{\frac{2\alpha+1}{3\alpha+1}+\eta})$ for any $\eta > 0$. $\square$

THEOREM 5.3. *Assume that for a certain learning scheme* $P(m^* - m(U_i^*) > \epsilon) = O(r_i)$ *for some known sequence* r_i. *Then for the certainty equivalence control with forcing-type scheme constructed above with the sequence* $\tau_i = \lfloor br_i^{-1}\rfloor$ *for some* $b > 0$, *we have* $E[L_n^\epsilon] = O(\kappa(n))$ *for all* $\epsilon > 0$.

Proof.

$$\begin{aligned} E[L_n^\epsilon] &= \sum_{l=1}^{n} P(m^* - m(U_l) > \epsilon) \\ &\leq \kappa(n) + \sum_{i=1}^{\kappa(n)} \sum_{l=t_i+1}^{t_{i+1}-1} P(m^* - m(U_l) > \epsilon) \\ &= \kappa(n) + \sum_{i=1}^{\kappa(n)} P(m^* - m(U_i^*) > \epsilon)\tau_i \\ &\leq \kappa(n) + \sum_{i=1}^{\kappa(n)} P(m^* - m(U_i^*) > \epsilon) br_i^{-1}. \end{aligned}$$

Now, we are given that $P(m^* - m(U_i^*) > \epsilon) = O(r_i)$. Thus, $\overline{\lim}_i P(m^* - m(U_i^*) > \epsilon)r_i^{-1} \leq C$ for some $C \geq 0$. That is, for all $\eta > 0$ there exists an $i_0 \geq 1$, such that for all $i \geq i_0$, $P(m^* - m(U_i^*) > \epsilon) \leq (C + \eta)r_i$. Then, clearly

$$\begin{aligned} E[L_n^\epsilon] &\leq \kappa(n) + \sum_{i=1}^{i_0-1} P(m^* - m(U_i^*) > \epsilon)br_i^{-1} + \sum_{i=i_0}^{\kappa(n)} P(m^* - m(U_i^*) > \epsilon)br_i^{-1} \\ &\leq \kappa(n) + M(\eta) + (C + \eta)b\kappa(n), \end{aligned}$$

where $M(\eta) = \sum_{i=1}^{i_0-1} P(m^* - m(U_i^*) > \epsilon)br_i^{-1}$ depends on η but not on n. Dividing by $\kappa(n)$ and taking the limit as $n \to \infty$ we get

$$\overline{\lim_n} \frac{E[L_n^\epsilon]}{\kappa(n)} \leq 1 + (C + \eta)b.$$

Now, by letting $\eta \to 0$ we get

$$\overline{\lim_n} \frac{E[L_n^\epsilon]}{\kappa(n)} \leq 1 + Cb. \qquad \square$$

COROLLARY 5.4. *For the certainty equivalence control with forcing-type scheme constructed above, with the learning scheme of* §3 *with any bandwidth sequence* h_i *such that* (i) $h_i \to 0$ *as* $i \to \infty$ *and* (ii) $ih_i \geq A'i^{a'}$ *for some* $0 < a' < 1$, $A' > 0$, *and with the sequence* $\tau_i = \lfloor be^{i(h_i)^2} \rfloor$ *for some* $b > 0$, *we have*

$$E[L_n^\epsilon] = O((h_n)^{-2} \log n), \tag{46}$$

for all $\epsilon > 0$.

Proof. This corollary will follow immediately from Theorems 5.3 and 3.1, if we can show that

$$\kappa(n) = O((h_n)^{-2} \log n) \tag{47}$$

when $\tau_i = \lfloor be^{i(h_i)^2} \rfloor$. Now by the definition of $\kappa(n)$, it follows that

$$\sum_{k=1}^{\kappa(n)-1} \tau_k + \kappa(n) \leq n.$$

Taking only the last term in the above sum we get

$$(\kappa(n) - 1) \leq h_{\kappa(n)-1}^{-2} (\log n - \log b) \leq h_n^{-2} (\log n - \log b).$$

This establishes (47). $\square$

Remark. In light of (46) above, we can choose $h_n \to 0$ arbitrarily slowly, thereby making $E[L_n^\epsilon]$ arbitrarily close to $O(\log n)$.

6. Concluding remarks. The ϵ-learning loss of the class of adaptive control schemes constructed in this paper for the case of an infinite number of arms is of the same order as those obtained previously for the finite case. For the finite case it is easy to see that the learning loss is within a constant factor of the ϵ-learning loss. The infinite case is fundamentally different in this respect. Thus, the learning loss that we obtain for the infinite case is considerably worse than those available for the finite case. Since we do not have any tighter lower bounds on the learning loss other than those available for the finite case, there may be room for improvement. However, to the best of our knowledge, these are the best rates available to date. Moreover, the rates obtained by us are still stronger than the $o(n)$ required for optimality with respect to the average-cost-per-unit-time criterion.

In a forthcoming paper [1] we extend the results of this paper to the multiarmed bandit problem as well as to the adaptive control of Markov chains, both with a control set $\mathcal{U}$ which is a bounded subset of $\mathbb{R}^d$, $d > 1$. The principal difficulty with this extension is that strong limit laws for moving averages in higher dimensions are not available in the literature, and are also much more difficult to obtain for the kind of sampling/design scheme that we employ.

Appendix: Proof of (15), (16), (27). First let $l(n) := \lfloor \log_2 n \rfloor + 1$, and consider the representation of n in base 2:

$$n = n_{l(n)} \dots n_2 n_1.$$

Note that we can partition $\{1 \dots n\}$ as

$$\{1 \dots n\} = \bigcup_{\substack{i=1,\dots,l(n):\\ n_i=1}} E_i^n,$$

where for $i = 1, \dots, l(n)$ such that $n_i = 1$, E_i^n are defined by

$$\begin{aligned} E_i^n &= \sum_{l=i+1}^{l(n)} n_l 2^{l-1} + \{1, \dots, 2^{i-1}\} \\ &= \{(n_{l(n)} \dots n_{i+1} 00 \dots 01), \dots, \\ &\qquad (n_{l(n)} \dots n_{i+1} 01 \dots 11), (n_{l(n)} \dots n_{i+1} 10 \dots 00)\} \\ &=: \{\underline{e}_i^n, \dots, \overline{e}_i^n\}. \end{aligned}$$

We can now partition $\{\tilde{u}_i\}_{i=1}^n$ correspondingly as

$$\{\tilde{u}_i\}_{i=1}^n = \bigcup_{\substack{i=1,\dots,l(n):\\ n_i=1}} \{\tilde{u}_k, k \in E_i^n\}.$$

It is easy to see that $\{\tilde{u}_k, k \in E_i^n\}$ is a uniform lattice of 2^{i-1} points in $[0, 1]$. More precisely,

$$\tilde{u}_k = \tilde{u}_{\underline{e}_i^n} + \tilde{u}_{k+1-\underline{e}_i^n}$$

for $k \in E_i^n$. As k ranges over the set E_i^n, the first term on the RHS stays fixed, and the second term gives us precisely the set of all dyadic rationals in $[0, 1]$ with denominators dividing 2^{i-1}. There are a total of 2^{i-1} such points. Thus we can think of $\{\tilde{u}_i\}_{i=1}^n$ as an overlay of these lattices of various levels of coarseness.

Then, we get for all $n \geq 1$ and $u \in \mathcal{U}$,

$$\begin{aligned} \sum_{i=1}^n I\{|u - \tilde{u}_i| \leq h_n/2\} &= \sum_{\substack{i=1,\dots,l(n):\\ n_i=1}} \sum_{k \in E_i^n} I\{|u - \tilde{u}_k| \leq h_n/2\} \\ &\leq \sum_{\substack{i=1,\dots,l(n):\\ n_i=1}} (2^{i-1} h_n + 1) \\ &= n h_n + l(n). \end{aligned}$$

Similarly, for all $n \geq n_1$ and $u \in \mathcal{U}$,

$$\begin{aligned} \sum_{i=1}^n I\{|u - \tilde{u}_i| \leq h_n/2\} &= \sum_{\substack{i=1,\dots,l(n):\\ n_i=1}} \sum_{k \in E_i^n} I\{|u - \tilde{u}_k| \leq h_n/2\} \\ &\geq \sum_{\substack{i=1,\dots,l(n):\\ n_i=1}} (2^{i-1} h_n - 1) \\ &= n h_n - l(n). \end{aligned}$$

This establishes (15).

Also, all $n \geq 1$ and $u \in \mathcal{U}$,

$$\begin{aligned}
&\sum_{i=1}^{n} I\{|u-\tilde{u}_i| \leq h_n/2\}|\tilde{u}_i - u|^\alpha \\
&= \sum_{\substack{i=1,\dots,l(n):\\ n_i=1}} \sum_{k \in E_i^n} I\{|u-\tilde{u}_k| \leq h_n/2\}|\tilde{u}_k - u|^\alpha \\
&= \sum_{\substack{i=1,\dots,l(n):\\ n_i=1}} \sum_{k \geq 0} [(s(u,n,i)+k2^{1-i})^\alpha I\{(s(u,n,i)+k2^{1-i}) \leq h_n/2\} \\
&\qquad +(2^{1-i}-s(u,n,i)+k2^{1-i})^\alpha I\{(2^{1-i}-s(u,n,i)+k2^{1-i}) \leq h_n/2\}],
\end{aligned}$$

where $0 \leq s(u,n,i) < 2^{1-i}$ is the distance from u to the closest point in $E_i^n \cap [u,1]$. Next, note that

$$\begin{aligned}
&\sum_{\substack{i=1,\dots,l(n):\\ n_i=1}} \sum_{k \geq 0} [(s(u,n,i)+k2^{1-i})^\alpha I\{(s(u,n,i)+k2^{1-i}) \leq h_n/2\} \\
&\qquad +(2^{1-i}-s(u,n,i)+k2^{1-i})^\alpha I\{(2^{1-i}-s(u,n,i)+k2^{1-i}) \leq h_n/2\}] \\
&\leq \sum_{\substack{i=1,\dots,l(n):\\ n_i=1}} \sum_{k \geq 0} 2\left[2^{i-1}\int_0^{h_n/2} x^\alpha dx + \left(\frac{h_n}{2}\right)^\alpha\right] \\
&= 2\sum_{\substack{i=1,\dots,l(n):\\ n_i=1}} \sum_{k \geq 0} \left[2^{i-1}\frac{1}{\alpha+1}\left(\frac{h_n}{2}\right)^{\alpha+1} + \left(\frac{h_n}{2}\right)^\alpha\right] \\
&= 2\left[\frac{1}{\alpha+1}\left(\frac{h_n}{2}\right)^{\alpha+1} n + \left(\frac{h_n}{2}\right)^\alpha l(n)\right] \\
&= \frac{1}{\alpha+1}\left(\frac{h_n}{2}\right)^\alpha [nh_n + 2(\alpha+1)l(n)],
\end{aligned}$$

which establishes (16). Similarly,

$$\begin{aligned}
&\sum_{\substack{i=1,\dots,l(n):\\ n_i=1}} \sum_{k \geq 0} [(s(u,n,i)+k2^{1-i})^\alpha I\{(s(u,n,i)+k2^{1-i}) \leq h_n/2\} \\
&\qquad +(2^{1-i}-s(u,n,i)+k2^{1-i})^\alpha I\{(2^{1-i}-s(u,n,i)+k2^{1-i}) \leq h_n/2\}] \\
&\geq \sum_{\substack{i=1,\dots,l(n):\\ n_i=1}} \sum_{k \geq 0} 2\left[2^{i-1}\int_0^{h_n/2} x^\alpha dx - \left(\frac{h_n}{2}\right)^\alpha\right] \\
&= \frac{1}{\alpha+1}\left(\frac{h_n}{2}\right)^\alpha [nh_n - 2(\alpha+1)l(n)],
\end{aligned}$$

which establishes (27).

Acknowledgments. The author would like to thank Professor J. Kuelbs for some very helpful discussions, and also two anonymous referees for comments that helped improve the exposition.

REFERENCES

[1] R. Agrawal, *Adaptive control of i.i.d. processes and Markov chains with a multidimensional control set*, working paper, Dept. of Electrical and Computer Engineering, Univ. of Wisconsin–Madison, 1992.

[2] R. Agrawal, M. Hegde, and D. Teneketzis, *Asymptotically efficient adaptive allocation rules for the multi-armed bandit problem with switching cost*, IEEE Trans. Automat. Control, AC-33 (1988), pp. 899–906.

[3] ———, *Multi-armed bandit problems with multiple plays and switching cost*, Stochastics Stochastic Rep., 29 (1990), pp. 437–459.

[4] R. Agrawal and D. Teneketzis, *Certainty equivalence control with forcing: Revisited*, Systems Control Lett., 13 (1989), pp. 405–412.

[5] R. Agrawal, D. Teneketzis, and V. Anantharam, *Asymptotically efficient adaptive allocation schemes for controlled i.i.d. processes: Finite parameter space*, IEEE Trans. Automat. Control, AC-34 (1989), pp. 258–267.

[6] V. Anantharam, P. Varaiya, and J. Walrand, *Asymptotically efficient allocation rules for the multiarmed bandit problem with multiple plays: Part* I: *IID rewards*, IEEE Trans. Automat. Control, 32 (1987), pp. 968–975.

[7] ———, *Asymptotically efficient allocation rules for the multiarmed bandit problem with multiple plays: Part* II: *Markovian rewards*, IEEE Trans. Automat. Control, 32 (1987), pp. 975–982.

[8] T. S. Chiang, C. R. Hwang, and S. J. Sheu, *Diffusions for global optimization in* $\mathbb{R}^n$, SIAM J. Control Optim., 25 (1987), pp. 737–752.

[9] M. Csörgő and P. Révész, *Strong Approximations in Probability and Statistics*, Academic Press, New York, 1981.

[10] A. de Acosta and J. Kuelbs, *Limit theorems for moving averages of independent random vectors*, Z. Wahrscheinlichkeitstheorie verw. Gebiete, 64 (1983), pp. 67–123.

[11] L. P. Devroye, *The uniform convergence of nearest neighbor regression function estimators and their application in optimization*, IEEE Trans. Inform. Theory, IT-24 (1978), pp. 142–151.

[12] V. Fabian, *Stochastic approximation*, in Optimizing Methods in Statistics, J. S. Rustagi, ed., Academic Press, New York, 1971, pp. 439–470.

[13] ———, *Polynomial estimation of regression functions with the supremum norm error*, Ann. Statist., 16 (1988), pp. 1345–1368.

[14] S. B. Gelfand and S. Mitter, *Recursive stochastic algorithms for global optimization in* $\mathbb{R}^d$, SIAM J. Control Optim., 29 (1991), pp. 999–1018.

[15] S. Geman and C. R. Hwang, *Diffusions for global optimization*, SIAM J. Control Optim., 24 (1986), pp. 1031–1043.

[16] W. Härdle, *Applied Nonparametric Regression*, Cambridge University Press, Cambridge, 1990.

[17] W. Härdle, P. Janssen, and R. Serfling, *Strong uniform consistency rates for estimators of conditional functionals*, Ann. Statist., 16 (1988), pp. 1428–1449.

[18] W. Härdle and S. Luckhaus, *Uniform consistency of a class of regression function estimators*, Ann. Statist., 12 (1984), pp. 612–623.

[19] J. Hoffman-Jørgensen, *Sums of independent Banach space valued random variables*, Studia Math., 52 (1974), pp. 159–186.

[20] H. J. Kushner, *Asymptotic global behavior for stochastic approximation and diffusions with slowly decreasing noise effects*, SIAM J. Appl. Math., 47 (1987), pp. 169–185.

[21] H. J. Kushner and D. S. Clark, *Stochastic Approximation Methods for Constrained and Unconstrained Systems*, Springer-Verlag, Berlin, New York, 1978.

[22] T. L. Lai and H. Robbins, *Asymptotically optimal allocation of treatments in sequential experiments*, in Design of Experiments, T. J. Santer and A. J. Tamhane, eds., Marcel Dekker, New York, 1984, pp. 127–142.

[23] ———, *Asymptotically efficient adaptive allocation rules*, Adv. in Appl. Math., 6 (1985), pp. 4–22.

[24] Y. P. Mack and B. W. Silverman, *Weak and strong uniform consistency of kernel regression estimates*, Z. Wahrscheinlichkeitstheorie verw. Gebiete, 61 (1982), pp. 405–415.

[25] M. B. Nevel'son and R. Z. Has'minskii, *Stochastic Approximation and Recursive Estimation*, American Mathematical Society, Providence, RI, 1973.

[26] H. Robbins, *Some aspects of sequential design of experiments*, Bull. Amer. Math. Soc., 55 (1952), pp. 527–535.

[27] C. J. Stone, *Optimal global rates of convergence for nonparametric regression*, Ann. Statist., 10 (1982), pp. 1040–1053.

[28] S. Yakowitz and W. Lowe, *Nonparametric bandit methods*, Ann. Oper. Res., 28 (1991), pp. 297–312.