Adaptive Hard Handoff Algorithms

Rajat Prakash, *Student Member, IEEE,* and Venugopal V. Veeravalli, *Senior Member, IEEE*

***Abstract*—The design of hard handoff algorithms based on optimizing the tradeoff between link quality and rate of handoffs is considered. For handoff algorithms based on this criterion, adaptation is precisely defined in terms of remaining on a locus of desirable operating points as system parameters (such as mobile velocity) change. A rule based on a linear cost criterion is used to select desirable operating points. For this rule, it is shown that the optimal handoff algorithm, which is impractical, is easily adapted by fixing a single tradeoff parameter at an appropriate value. The same adaptation property is shown to hold for an easily implementable approximation to the optimal algorithm, the locally optimal (LO) handoff algorithm. This is in contrast to the poor adaptation of hysteresis based approaches which require lookup tables for adaptation. Practical estimators for all relevant system parameters based on a short window of pilot signal strength measurements are also discussed. It is shown that the LO algorithm adapts well when these simple estimators are used. A hysteresis-threshold approximation to the adaptive LO algorithm is also developed.**

***Index Terms*—Adaptive systems, cellular land mobile radio, handover.**

I. Introduction

A HANDOFF in cellular communications is the process whereby a mobile subscriber communicating with one base station is switched to another base station during a call. Handoff algorithm design is an important component of the larger problem of radio resource management in mobile communication systems. In this paper we concentrate on mobile-assisted hard handoffs in a channelized cellular system. The term handoff is henceforth used to refer to hard handoff.

The decision to handoff from one cell to the other is based on various criteria that take into account channel degradation considerations, as well as Erlang capacity and blocking considerations. However, the initial (and most important) trigger for a handoff is generally based on pilot signal strength measurements[1] taken at the mobile [1]. We focus on this initial trigger event and assume that a handoff occurs if and only if such a trigger is present. Thus, in our terminology, a handoff event is the occurrence of the initial trigger.

As argued in [2], the handoff algorithm design problem may be posed as an optimization to obtain the best tradeoff between the expected rate of handoffs λ_{H} and the link quality. The measure of link quality is the expected rate of link degradation events λ_{LD} on a mobile trajectory. A link degradation event corresponds to the pilot power level falling below a threshold Δ. Under the assumptions that 1) the peak power of the transmitters is limited, 2) the link gains seen by the pilot and traffic channels are the same, and 3) the interference level is known and does not change significantly with time, it is easy to see that a measure of traffic channel quality may be derived from the rate of link degradation events. However, if the interference level varies with time, and can be tracked by other sensors, this information can be used to improve the handoff algorithm by varying the threshold Δ.

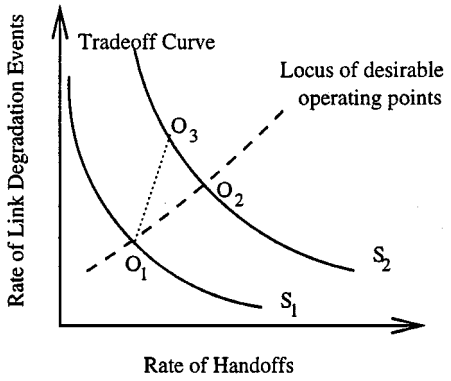

Fig. 1. Tradeoff curves on $\lambda_{\mathrm{LD}} - \lambda_{\mathrm{H}}$ plane.

The optimization to obtain the best tradeoff between λ_{LD} and λ_{H} can be solved using dynamic programming (DP) [3] to yield an optimal handoff algorithm (see, e.g., [2]). A locally optimal (LO) handoff algorithm was introduced in [2] as a practical approximation to the optimal algorithm. It was conjectured in [2] that the LO algorithm can easily be made to adapt to changes in system parameters. This adaptation property is established precisely in this paper.

For a general handoff algorithm, let ξ denote the handoff algorithm parameters, and let $\mathcal{S}$ denote the vector of system parameters. The vector $\mathcal{S}$ may include mobile velocity and propagation characteristics, while ξ may be the hysteresis level for the case of a simple hysteresis handoff algorithm. Given a fixed choice of system parameters, say $\mathcal{S}_1$, varying ξ over all choices of handoff algorithm parameters results in a set of feasible operating points $(\lambda_{\mathrm{LD}}(\xi), \lambda_{\mathrm{H}}(\xi))$ on the $\lambda_{\mathrm{LD}} - \lambda_{\mathrm{H}}$ plane (Fig. 1). If the handoff parameter ξ is unidimensional, this set of feasible operating points is defined to be the tradeoff curve for the system parameters $\mathcal{S}_1$ and the given handoff algorithm.

In case ξ is multidimensional, the set of feasible operating points is a region on the $\lambda_{\mathrm{LD}} - \lambda_{\mathrm{H}}$ plane. The lower envelope of this region defines the best tradeoff curve for the given $\mathcal{S}_1$ and the given handoff algorithm. Alternatively, a tradeoff curve

Manuscript received March 17, 1999; revised March 11, 2000. This work was supported in part by NSF CAREER/PECASE under Grant CCR-9733204, by the Office of Naval Research under Grant N0014-97-1-0823, and by Nortel Networks' Global External Research. This work was presented in part at the IEEE Vehicular Technology Conference, Houston, TX, 1999.

The authors are with the ECE Department and the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, IL 61801 USA (e-mail: {rprakasl,vvv}@uiuc.edu).

Publisher Item Identifier S 0733-8716(00)09193-9.

[1]Often, pilot signal strength measurements are the only measurements available for the candidate handoff base stations, because no traffic channel link exists between these base stations and the mobile.

may be generated by varying only one of the parameters ξ, while fixing the rest.

In order to select a suitable ξ for the given $\mathcal{S}_1$, a rule is needed to select a desirable operating point on the tradeoff curve corresponding to $\mathcal{S}_1$. This rule may be based on higher layer considerations, and may be in the form of a constraint curve on the $\lambda_{\mathrm{LD}} - \lambda_{\mathrm{H}}$ plane. The desirable operating point, say, O_1, is then defined as the intersection of the tradeoff curve and the constraint curve. Alternatively, the desirable operating point may be defined as the point on the tradeoff curve where a given cost metric, which may be imposed by higher layer considerations, is minimized. Let the value of the handoff algorithm parameters at the desirable operating point be ξ_1.

Changes in $\mathcal{S}$ may cause the tradeoff curve to shift (see Fig. 1). Let the new system parameters be $\mathcal{S}_2$, and let the desirable operating point on the new tradeoff curve be O_2. The old choice of algorithm parameters ξ_1 may result in an operating point $O_3 \neq O_2$. *Adaptation* of a handoff algorithm involves estimation of $\mathcal{S}_2$, and finding $\xi_2 = f(\mathcal{S}_2)$ to yield a desirable operating point on the tradeoff curve corresponding to $\mathcal{S}_2$.

It is important that the function f as well as the estimators of $\mathcal{S}$ be easy to implement. Although every function f can be approximated by a lookup table, construction of this lookup table would require obtaining tradeoff curves for different values of $\mathcal{S}$ through simulations or repeated field trials. For multidimensional $\mathcal{S}$, the lookup tables also need large memory for storage. We show that for the LO algorithm, the adaptation function can be obtained in closed form, and does not require the construction of lookup tables.

If a rule for the selection of desirable operating points is not given, an *ad hoc* criterion such as trying to remain on the "knee" of the tradeoff curve may be adopted. This ill-defined criterion has been used in previous papers on adaptive hard handoff algorithms. We develop a more systematic rule, based on a general linear cost structure, and use it to define adaptation.

Previous work on adaptive hard handoff algorithms [4]–[6] has considered the tradeoff between number of handoffs and handoff delay. Desirable operating points are defined using the "knee" criterion in these papers. Also considered in [5] and [6] is the estimation of the changing system parameters. Estimation of shadow fading variance σ^2 as well as adaptation of hysteresis level and averaging interval length to changes in σ are considered in [5]. Estimators of velocity are considered in [6], where averaging interval length is adapted as a function of the velocity estimate.

Our approach differs from previous work in the following ways. First, we give a precise rule for defining the locus of desirable operating points based on a general linear cost structure. Second, we substitute handoff delay with a performance metric that is directly related to link quality. Third, the estimates of the system parameters that we use are based solely on pilot signal strength measurements. No other information, such as Doppler measurements for estimating velocity, or distance information as in [7], is used in our approach. Finally, and most importantly, we derive adaptive handoff algorithms for which the adaptation rule $\xi = f(\mathcal{S})$ is explicit and does not require the construction of lookup tables. Furthermore, this function can be obtained by using a single field-trial or simulation for a typical set of system parameters values.

This paper is organized as follows. In Section II a mathematical structure for the handoff problem is developed. The adaptation property and the optimal handoff algorithm are described in Section III. Adaptation of the LO algorithm is considered in Section IV. In Section V, an adaptive hysteresis-threshold approximation of the LO algorithm is developed, and results on its performance are presented. Conclusions are given in Section VI. Finally, estimators of system parameters are presented in the Appendix.

II. Preliminaries

Assume that only two base stations (BS's), say BS-1 and BS-2, are involved in the handoff, i.e., consider only that portion of the trajectory on which the signals received from these two base stations are the strongest. The extension to the case where more than two base stations are involved is not difficult [8]. The analysis is restricted to short time horizons over which it may be assumed that the mobile is moving on a straight line with fixed velocity v. Over the time horizon of interest, let N samples of received pilot power be taken at equally spaced time intervals t_s seconds apart. Let $X_{k,1}$ and $X_{k,2}$ be the received pilot signal strengths at the kth sampling instant (time kt_s) from BS-1 and BS-2, respectively. Also, let the respective distances of the mobile from BS-1 and BS-2 be $d_{k,1}$ and $d_{k,2}$. Now assume the following propagation model for base station i

$$X_{k,i} = \overline{P}_{k,i} + Z_{k,i} \quad \text{dBm} \tag{1}$$

$$\overline{P}_{k,i} = \mu_i - 10\eta_i \log d_{k,i} \tag{2}$$

where $\overline{P}_{k,i}$ is the local mean pilot power, and $Z_{k,i}$ is the shadow fading at the kth sampling instant. Short-term variations due to multipath are assumed to be either averaged out or combatted through diversity. A log-normal first order autoregressive (AR-1) model is assumed for shadow fading [9]. Therefore, under the assumption of equally spaced time samples and constant v, $\{Z_{k,i}\}$ is a zero mean AR-1 Gaussian process with autocorrelation function,

$$\mathrm{E}[Z_{k,i} Z_{k+m,i}] = \sigma_i^2 a_i^{|m|}.$$

See Fig. 2 for a typical propagation characteristic. Furthermore, $\{Z_{k,i}\}$ can be written recursively as

$$\begin{aligned} Z_{0,i} &= \sigma_i^2 W_{0,i} \\ Z_{k+1,i} &= a_i Z_{k,i} + \sigma_i \sqrt{1 - a_i^2} W_{k,i} \end{aligned} \tag{3}$$

where

- $\{W_{k,i}\}$ are independent and identically distributed (i.i.d.) $\mathcal{N}(0, 1)$ random variables,
- $\overline{d}_i$ is the shadow fading correlation distance,
- σ_i^2 is the shadow fading variance, and
- a_i is the correlation coefficient of $\{Z_{k,i}\}$, i.e.,

$$a_i = \exp(-vt_s/\overline{d}_i). \tag{4}$$

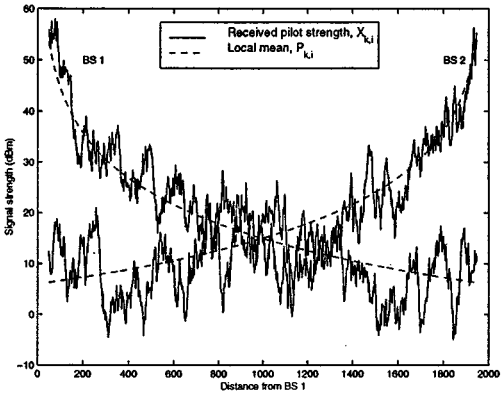

Fig. 2. Typical profile of received pilot power from two base stations as mobile moves along the line joining the base stations. The solid line is the local mean received power component which decreases monotonically as the mobile moves away from the base station. The dotted line includes shadow fading. Parameters are as given in Table I.

Under this model, the distribution of $X_{k+1,i}$ conditioned on $X_{k,i}$, is independent of earlier received power samples (a standard result for AR-1 Gaussian processes), and is described completely by its conditional mean and variance

$$\begin{aligned}\mathrm{E}[X_{k+1,i}|X_{k,1}\cdots X_{k,i}] &= \mathrm{E}[X_{k+1,i}|X_{k,i}] \\ &= \overline{P}_{k+1,i} + a_i(X_{k,i} - \overline{P}_{k,i}) \\ \mathrm{Var}[X_{k+1,i}|X_{k,1}\cdots X_{k,i}] &= \mathrm{Var}[X_{k+1,i}|X_{k,i}] \\ &= (1-a_i^2)\sigma_i^2. \end{aligned} \tag{5}$$

In the current model, there is no correlation between the shadow fading from the two base stations, although the analysis of this paper extends easily to the case where the fading is correlated [8].

Let B_k denote the index of the operative base station at sampling instant k (i.e., $B_k = i$, when the mobile is communicating with base station i), and let $\overline{B}_k$ denote the other base station. We assume that a handoff decision is made during each sampling interval. The decision variable U_k takes on two values. If $U_k = 1$, a handoff is made, resulting in $B_{k+1} = \overline{B}_k$; if $U_k = 0$, no handoff is made, and $B_{k+1} = B_k$. Define the state vector S_k as

$$S_k = (X_{k,1},\, X_{k,2},\, B_k). \tag{6}$$

Let I_k denote all the information available for decision making at time k, i.e.,

$$I_k = (S_k,\, S_{k-1},\, \cdots S_1). \tag{7}$$

The decision U_k at time k can be based on all signal strength measurements up to time k, i.e., $U_k = \phi_k(I_k)$, where ϕ_k is the decision function at time k. The sequence of decision functions, $\Phi = \{\phi_k(\cdot)\}$, constitutes the handoff policy.

An example of a handoff algorithm is the simple hysteresis algorithm [1]. The handoff policy Φ of the hysteresis algorithm, is described by the following equation:

$$X_{k,\overline{B}_k} \underset{U_k=0}{\overset{U_k=1}{\gtrless}} X_{k,B_k} + h \tag{8}$$

i.e., a handoff occurs whenever the other base station's power exceeds the current base station's power by a level h. The handoff algorithm parameter $\xi = h$ completely defines the handoff policy $\Phi = \Phi_{\mathrm{Hy}}(\xi)$ for the hysteresis algorithm.

We measure the performance of a handoff algorithm in terms of the expected rate of handoffs (λ_{H}) and the expected rate of link degradation events (λ_{LD}), the rates being defined per unit time by

$$\begin{aligned}\lambda_{\mathrm{LD}}(\Phi) &= \mathrm{E}\left[\frac{1}{Nt_s}\sum_{k=1}^{N} \mathbb{I}_{\{X_{k,B_k} < \Delta_{B_k}\}}\right] \\ \lambda_{\mathrm{H}}(\Phi) &= \mathrm{E}\left[\frac{1}{Nt_s}\sum_{k=1}^{N} \mathbb{I}_{\{U_k=1\}}\right] \end{aligned} \tag{9}$$

where $\mathbb{I}_{\{\cdot\}}$ is the indicator function and the expectation is carried out over all realizations of the state sequence $S_1, \cdots, S_N$, given the handoff algorithm Φ. The average is taken over N samples, where Nt_s is the duration of the call. The parameter Δ_i is the minimum pilot signal strength required for satisfactory communication with BS-i, and the pilot strength falling below Δ_i results in a link degradation event. The performance metrics λ_{LD} and λ_{H} are clearly a function of the handoff policy Φ, and hence we write them explicitly as $\lambda_{\mathrm{LD}}(\Phi)$ and $\lambda_{\mathrm{H}}(\Phi)$. Further, for a given family of handoff algorithms with handoff parameters ξ, with slight abuse of notation, we can write the performance metrics as $\lambda_{\mathrm{LD}}(\xi)$ and $\lambda_{\mathrm{H}}(\xi)$.

III. Adaptation and the Optimal Handoff Algorithm

To determine a desirable operating point on the tradeoff curve corresponding to a given handoff algorithm and given $\mathcal{S}$, we use a linear cost criterion. That is, we assume that the goal is to minimize a linear combination of the performance metrics $\lambda_{\mathrm{LD}}(\xi)$ and $\lambda_{\mathrm{H}}(\xi)$

$$J(\xi) = \lambda_{\mathrm{LD}}(\xi) + \gamma\lambda_{\mathrm{H}}(\xi), \qquad \text{for a given } \gamma > 0 \tag{10}$$

over the set of feasible ξ. Note that any general linear cost function can be written in the form given in (10) for the purposes of minimization. Nonlinear cost criteria are discussed at the end of this section.

We assume that the tradeoff curve is convex. This is a reasonable assumption because convex tradeoff curves are obtained for the optimal, LO, hysteresis and hysteresis-threshold algorithms in [2], and also for the algorithms considered in [4]–[6]. The convexity assumption can be extended to a general handoff algorithm (with more than one parameter) if randomization is allowed and the tradeoff curve is defined to be the lower envelope of all feasible operating points. If the tradeoff curve is differentiable everywhere, the cost function (10) is minimized by setting

$$\frac{\partial\lambda_{\mathrm{LD}}}{\partial\lambda_{\mathrm{H}}} + \gamma = 0. \tag{11}$$

Remark 1: If (11) has no solution, the cost (10) is minimized at one of the end points of the tradeoff curve. If the assumption about the gradient of the tradeoff curve being continuous is relaxed, (10) is minimized at a point on the tradeoff curve where

$-\gamma$ lies between the value of the left derivative and the value of the right derivative at that point.

Using (10), the desirable operating point may be defined in terms of the gradient as given below.

Definition 1: A desirable operating point on a tradeoff curve is the point at which the ratio of the incremental change in $\lambda_{\rm LD}$ and the incremental change in λ_H, i.e., the gradient of the tradeoff curve, equals $-\gamma$.

This definition of desirable operating point is much more precise than the "knee" criterion used in previous work on adaptive hard handoff algorithms [4]–[6], and we use it to define adaptation as follows.

Definition 2: A handoff scheme is said to be adaptive if it meets the following criteria.

1) It allows for the estimation of all relevant system parameters which can change during a call, using an appropriate estimator $\hat{\mathcal{S}}$.
2) The handoff algorithm parameters that are required to operate on a desirable operating point on the tradeoff curve can be written as an explicit function of $\hat{S}$.

We now consider adaptation of the optimal handoff algorithm. The optimal handoff algorithm attains the best tradeoff between the performance metrics by minimizing $\lambda_{\rm H}$ given a constraint on $\lambda_{\rm LD}$. It has been proved [2] that this optimum tradeoff problem can be posed in a Bayesian framework, with the optimum decision policy being given by

$$\Phi^*(c) = \arg\min_{\Phi}(\lambda_{\rm LD}(\Phi) + c\lambda_{\rm H}(\Phi)) \tag{12}$$

where the minimization is over all possible Φ. The tradeoff parameter c is the handoff algorithm parameter ξ for the class of optimal handoff algorithms. The parameter c can also be thought of as the relative cost of a handoff versus that of a link degradation event. Dynamic programming (DP) can be used to solve (12) recursively as shown in [2] and [10]. The locus of operating points of the optimal algorithm [i.e., the set of points $\lambda_{\rm LD}(\Phi^*(c))$, $\lambda_{\rm H}(\Phi^*(c))$] obtained by varying c is the tradeoff curve of the optimal algorithm. Tradeoff curves for for different mobile velocities are shown in Fig. 3.

We next show that for the optimal algorithm, the desirable operating point on the tradeoff curve for any system parameter value can be attained simply by setting $c = \gamma$. From [2], we know that the tradeoff curve for the optimal algorithm is convex. Thus, the point on the tradeoff curve where the gradient is $-\gamma$ is the same as the point on the tradeoff curve where $\lambda_{\rm LD} + \gamma\lambda_{\rm H}$ is minimized. By definition of the optimal algorithm, the above minimum is attained by solving the Bayes problem (12) for $c = \gamma$. Thus, fixing the algorithm parameter $c = \gamma$ guarantees operation on the locus of desirable operating points for all system parameter values. This result is verified for changing v in Fig. 3.

The optimal algorithm also has the obvious advantage of resulting in minimum cost among all handoff algorithms. However, the optimal algorithm is impractical because it requires prior knowledge of the entire trajectory and system parameters [2]. Also, while the adaptation of the optimal algorithm can be written explicitly as a function of system parameter estimates $\hat{\mathcal{S}}$ for the purposes of adaptation (see Definition 2), this function is complicated and not easily evaluated.

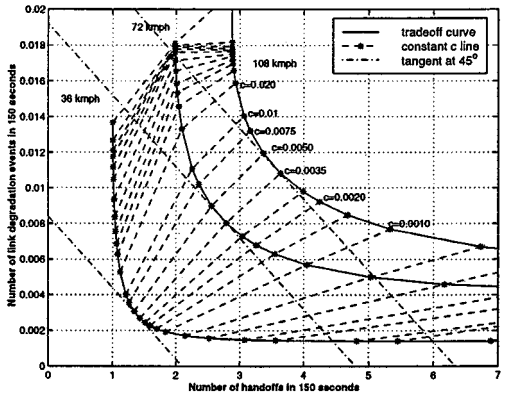

Fig. 3. Tradeoff curves for optimal algorithm with tangent lines at operating points with fixed slope. The tradeoff curves move up when velocity increases because the mobile covers a greater distance in the same time interval (causing the $\lambda_{\rm H}$ to increase). Prediction of signal strength at the next time instant also becomes difficult at higher speeds, resulting in higher $\lambda_{\rm LD}$.

We now comment on the case where a nonlinear cost structure is imposed by higher layers. The current approach can be applied to the nonlinear case by using standard minimization tools to find a desirable operating point for one set of system parameter values, and then linearizing the given cost around that point to obtain γ for an approximate linear cost structure as in (10). The usefulness of such an approximation will of course depend on the nature of the cost.

IV. Adaptation of the LO Algorithm

The LO handoff algorithm was suggested in [2] as a practical approximation to the optimal algorithm (Section III), and it also has c as a handoff algorithm parameter. The LO algorithm is essentially equivalent to the dynamic programming solution, with the time horizon limited to the immediate future sampling instant, $(k+1)$. The set of decision functions for the LO algorithm is governed by,

$$P\left\{X_{k+1,\,\overline{B}_k} < \Delta_{\overline{B}_k} | I_k\right\} + c \underset{U_k=1}{\overset{U_k=0}{\gtrless}} P\left\{X_{k+1,\,B_k} < \Delta_{B_k} | I_k\right\}. \tag{13}$$

Using (5), the above probabilities can be written in terms of the Q function, and the LO algorithm takes the form [2]

$$Q\left(\frac{\mathrm{E}[X_{k+1,\overline{B}_k}|I_k] - \Delta_{\overline{B}_k}}{\sigma_{\overline{B}_k}\sqrt{1-a^2_{\overline{B}_k}}}\right) + c \underset{U_k=1}{\overset{U_k=0}{\lessgtr}} Q\left(\frac{\mathrm{E}[X_{k+1,B_k}|I_k] - \Delta_{B_k}}{\sigma_{B_k}\sqrt{1-a^2_{B_k}}}\right). \tag{14}$$

To evaluate the above decision rule, σ_i, a_i and $\mathrm{E}[X_{k+1,i}|I_k]$ need to be estimated for both base stations ($i = 1, 2$) using estimators $\hat{\sigma}_i$, $\hat{a}_i$ and $\hat{X}_{k+1,i}$, respectively. The decision rule using these estimators is

$$Q\left(\frac{\hat{X}_{k+1,\overline{B}_k} - \Delta_{\overline{B}_k}}{\hat{\sigma}_{\overline{B}_k}\sqrt{1-\hat{a}^2_{\overline{B}_k}}}\right) + c \underset{U_k=1}{\overset{U_k=0}{\gtrless}} Q\left(\frac{\hat{X}_{k+1,B_k} - \Delta_{B_k}}{\hat{\sigma}_{B_k}\sqrt{1-\hat{a}^2_{B_k}}}\right). \tag{15}$$

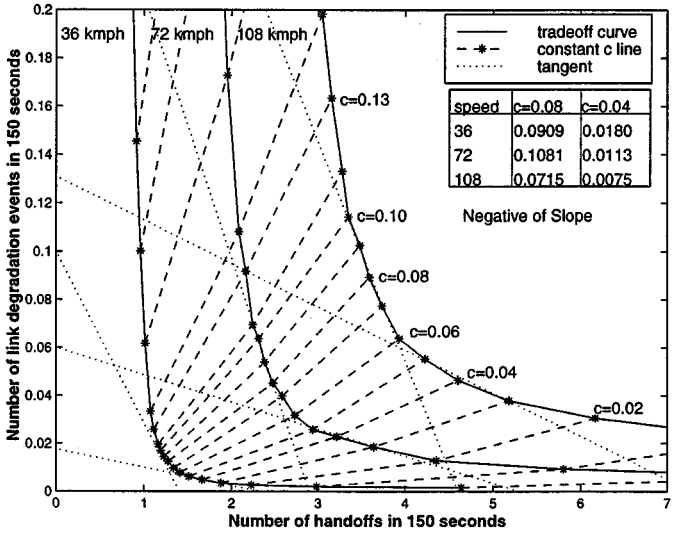

Fig. 4. Performance of the adaptive LO algorithm.

Examples of estimators based only on received pilot measurements are discussed in the Appendix. For the simulations in this paper, estimators $\hat{a}_i^{\mathrm{J}}$ and $\hat{\sigma}_i^{\mathrm{J}}$ are used to estimate the constant but unknown parameters a_i and σ_i. Estimator $\hat{X}_{k+1}^{\mathrm{S}}$ is used to estimate $\mathrm{E}[X_{k+1,i}|I_k]$. Among the estimators discussed in the Appendix, the ones selected for simulation are not the most precise ones, but rather the ones which are simple to implement. It is shown that the adaptation property of the LO algorithm holds even with these relatively weak estimators.

The LO algorithm has the following advantages over other algorithms. First, as shown in [2], the LO algorithm's performance is close to that of the optimal algorithm. Second, it is clear from (15) that the structure of the LO algorithm depends directly on the system parameter estimates. Thus, in contrast to the optimal and hysteresis-based algorithms, the LO algorithm can easily be adapted as the system parameters change. Finally, since the LO algorithm is derived as an approximation to the optimal algorithm, we expect it to inherit the adaptation property of the optimal algorithm, i.e., the slope at the operating point should remain fixed once c has been fixed.

Although the LO algorithm is expected to adapt to all system parameters, we use simulations to verify adaptation only to changes in v. Note that v can change significantly over a few meters, whereas parameters σ and $\overline{d}$ can generally be assumed to be constant over a large region of the cell. Changes in v influence the system parameter $a = \exp(vt_s/\overline{d})$ of the LO algorithm. Estimating a eliminates the need to separately estimate $\overline{d}$.

Fig. 4 shows the results of simulation for parameters values shown in Table I. It can be seen that fixing c indeed gives operating points with nearly the same gradient[2] for different v. Using Definition 1, it follows that operation at or near the desirable operating point can be maintained simply by fixing $c = c_\gamma$ appropriately. In other words, the adaptation function (Definition 2, part 2) for the LO algorithm is simply a constant, $f(\mathcal{S}) = c_\gamma$.

[2]Slight changes in gradient for fixed c are caused by imperfect estimators and also because the LO algorithm is an approximation to the optimal.

TABLE I
PARAMETERS USED FOR ALL THE SIMULATIONS. THE RESULTING FADE MARGIN IS 3σ

D= 2000 meters,	distance from BS-1 to BS-2
μ_i= 105 dBm,	station strength
η_i= 3,	path loss exponent
σ_i= 5 dBm,	shadow fading std. dev.
v= 10, 20 & 30 m/s,	mobile velocity
$\bar{d}$= 30 meters,	correlation distance
Δ_i= 0 dBm,	threshold of link degradation
t_s= 0.5s,	Sample Time
M= 20,	Window length for estimation

Recall that for the LO algorithm, the only handoff parameter ξ is the tradeoff parameter c.

The following procedure is used to select c_γ for the given γ of the linear cost function (10). Through field trials or simulation for a typical combination of system parameters $\mathcal{S}$, we can obtain a tradeoff curve, and select a value of c_γ corresponding to the desired gradient $-\gamma$. For $c = c_\gamma$, we can claim from the adaptation property that the operating point will be at a point with gradient close to $-\gamma$ even when $\mathcal{S}$ changes. Thus, the use of this procedure, along with estimators $\hat{X}_{k+1}^{\mathrm{S}}$, $\hat{a}^{\mathrm{J}}$ and σ^{J}, results in a simple and practical adaptive locally optimal handoff algorithm.

To stress the important role of estimators in the adaptation of the LO algorithm, we consider its performance with the estimators fixed at one value even as actual system parameters change. In particular, even as v changes, $\hat{a}$ is kept fixed to the value corresponding to $v = 36$ km/h. However, estimators $\hat{\sigma}^{\mathrm{J}}$ and $\hat{X}^{\mathrm{S}}$ are allowed to vary depending on the pilot measurements. Tradeoff curves for different velocities are shown in Fig. 5. Fixing c yields operating points with widely varying gradients, demonstrating that system parameter estimation is crucial to the adaptation of the LO algorithm.

The only disadvantage of the LO algorithm relative to the hysteresis based algorithm is the need to evaluate the Q function at each sample instant; this can be avoided by using the approximation technique given in the following section.

V. HYSTERESIS-THRESHOLD APPROXIMATION

An *ad hoc* handoff algorithm that is currently in use is the simple hysteresis algorithm (8). For this algorithm, however, there are no clear rules for selecting a suitable value of h, and its performance was shown to be much inferior to the LO algorithm in [2].

Performance of the hysteresis algorithm can be improved by incorporating a threshold on signal strength. The hysteresis-threshold algorithm was introduced in [11], and it can be generalized by including a threshold on the signal from the base station $\overline{B}_k$

$$U_k = \begin{cases} 1 & (\hat{X}_{k+1,\,\overline{B}_k} > t_1)\,\&\,(\hat{X}_{k+1,\,B_k} < t_2) \\ & \quad \&\,(\hat{X}_{k+1,\,\overline{B}_k} > \hat{X}_{k+1,\,B_k} + h) \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

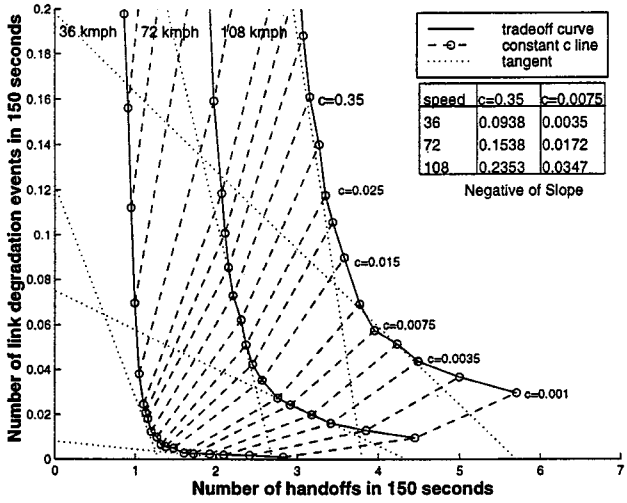

Fig. 5. Performance of a nonadaptive LO algorithm, with correlation coefficient a fixed to the value corresponding to speed 36 km/h.

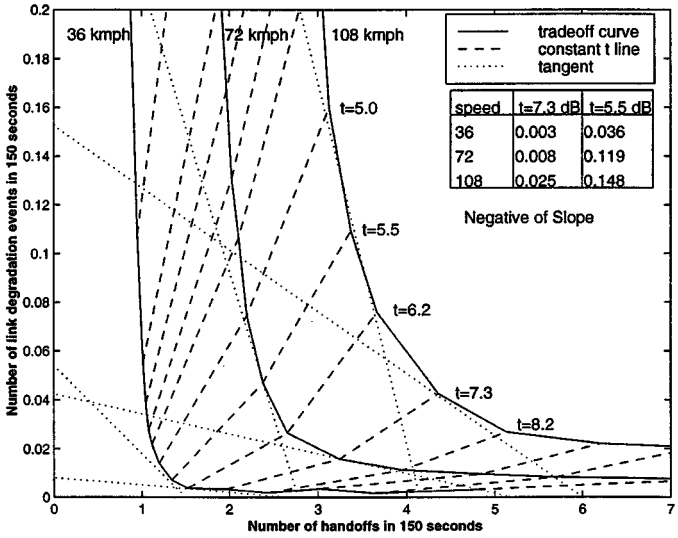

Fig. 6. Performance of hysteresis-threshold algorithm with hysteresis level fixed to 3 dB.

For the hysteresis-threshold algorithm also, there is no obvious way of selecting suitable values of quantities h, t_1, and t_2. Obtaining the best tradeoff curve (see Section I) for this algorithm is difficult because it is not clear which combination of the three parameters will result in the best tradeoff curve. Therefore, as in [2], we consider tradeoff curves obtained by fixing h, and varying t_1 ($t_2 = t_1$ is used). This hysteresis-threshold algorithm effectively has only one handoff algorithm parameter, $\xi = t_1$, and it results in a tradeoff curve, rather than a region (see Section I).

The rule for adapting ξ (Definition 2, part 2) as a function of system parameters does not follow naturally from the structure of the handoff algorithm (16). We now show that such a rule is important for the hysteresis-threshold algorithm. Tradeoff curves for a hysteresis-threshold algorithm, where ξ is not altered as a function of $\mathcal{S}$, are shown in Fig. 6. It can be seen that there is a considerable change in the slope at the operating points as velocity changes and ξ is kept fixed. Fig. 6 also reveals that a lookup table may be used to select ξ as a function of v in order to stay on an operating point with the desired gradient. (As mentioned in Section I, constructing such lookup tables requires extensive field trials or simulations.)

To develop simpler rules for adaptation of h, t_1, and t_2 to changes in system parameters we derive the hysteresis-threshold algorithm as an approximation to the LO algorithm. This approximation can also be considered as a way of implementing the LO algorithm without evaluating a Q function at each sample instant.

The decision function of the LO algorithm depends only on c and $\hat{\mathcal{S}}$ (17). To achieve the desired approximation, $(h, t_1, t_2) = g(c, \hat{\mathcal{S}})$ is derived to make the decision function of the hysteresis-threshold algorithm (16) close to the decision function of the LO algorithm (15). Since rules for adapting the LO algorithm as a function of $\mathcal{S}$ are known, the hysteresis-threshold algorithm derived in this way is expected to inherit the adaptation properties of the LO algorithm.

To derive the approximation, consider the handoff region of the LO algorithm. The handoff region at sampling instant k corresponds to the area on the $\hat{X}_{k+1,1}$, $\hat{X}_{k+1,2}$ plane where $U_k = 1$. Consider the case of handoff from BS-2 to BS-1. Let $\hat{\sigma}\sqrt{1-(\hat{a}^S)^2} = \sigma'$ be the same for both base stations (the derivation can be carried out without this assumption also). Then, the boundary of the handoff region for the LO algorithm (15) is given by

$$Q\left(\frac{\hat{X}_{k+1,1}-\Delta}{\sigma'}\right)+c = Q\left(\frac{\hat{X}_{k+1,2}-\Delta}{\sigma'}\right). \tag{17}$$

The handoff region and its boundary are shown in Fig. 7. Consider approximating this boundary (17) by three straight lines ℓ_1, ℓ_2, and ℓ_3 as shown in Fig. 7. Lines ℓ_1 and ℓ_2 correspond to the threshold parameters t_1 and t_2 in the definition of the hysteresis-threshold algorithm (16), while the line ℓ_3 corresponds to the hysteresis parameter h.

To find the equation of ℓ_1, set $\hat{X}_{k+1,2}$ to $-\infty$ in (17). Solving for $\hat{X}_{k+1,1}$ gives the asymptote

$$\hat{X}_{k+1,1} = \Delta - \sigma' Q^{-1}(c) \tag{18}$$

corresponding to the condition, $\hat{X}_{k+1,1} > t_1$ in (16). Comparing (16) and (18), we get

$$t_1 = \Delta - \sigma' Q^{-1}(c). \tag{19}$$

Similarly

$$t_2 = \Delta + \sigma' Q^{-1}(c). \tag{20}$$

To find the equation of ℓ_3, use the intercepts of (17) with the lines $\hat{X}_{k+1,1} = \Delta$ and $\hat{X}_{k+1,2} = \Delta$. Setting $\hat{X}_{k+1,1} = \Delta$ in (17) gives $\hat{X}_{k+1,2} = \Delta + \sigma' Q^{-1}(1/2 + c)$. Similarly, setting $\hat{X}_{k+1,2} = \Delta$ gives $\hat{X}_{k+1,1} = \Delta + \sigma' Q^{-1}(1/2 - c)$. Thus, the intercepts are at the points

$$(\Delta,\ \Delta + \sigma' Q^{-1}(1/2+c)) \quad \text{and} \quad (\Delta - \sigma' Q^{-1}(1/2+c), \Delta).$$

Joining these points gives the following equation for ℓ_3:

$$\hat{X}_{k+1,1} - \hat{X}_{k+1,2} = -\sigma' Q^{-1}(1/2+c).$$

Comparing to (16) gives

$$h = -\sigma' Q^{-1}(1/2+c). \tag{21}$$

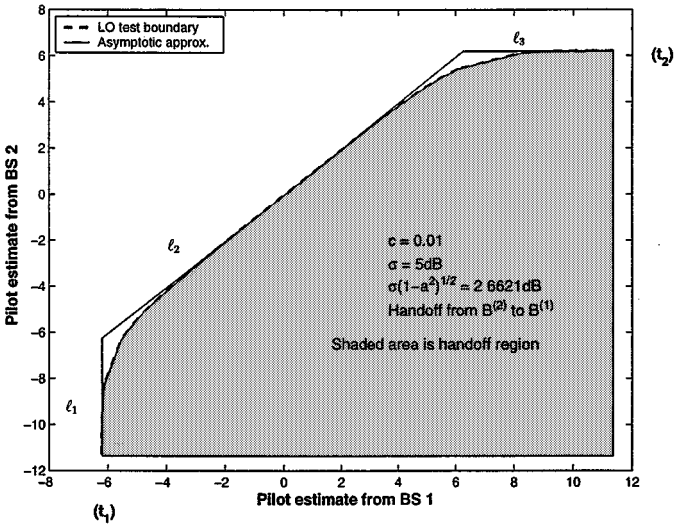

Fig. 7. Decision regions of the LO algorithm and its hysteresis-threshold approximation.

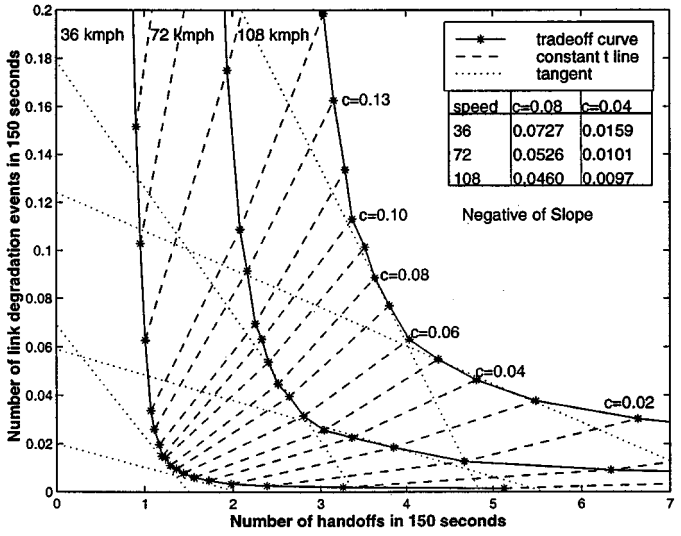

Fig. 8. Performance of the hysteresis-threshold approximation to the LO algorithm.

Thus, the LO algorithm is approximated by a hysteresis-threshold algorithm through the function g, described by (19)–(21).

As system parameters a and σ change, the hysteresis-threshold algorithm parameters can be adapted by using $\sigma' = \hat{\sigma}\sqrt{(1-(\hat{a}^J)^2)}$ in (19)–(21). Performance of this adaptive approximation is shown in Fig. 8. It can be seen that the adaptation property of the LO algorithm is preserved, i.e., the slope at an operating point remains nearly fixed once c is fixed.

It was seen in Section IV that for the purposes of adaptation, c for the LO algorithm is fixed at c_γ. Therefore, the inverse Q functions in (19)–(21) need not be computed at each sample instant. In order to adapt t_1, t_2, and h, only σ' needs to be be evaluated at each sample instant using the estimators $\hat{a}^S$ and $\hat{\sigma}^J$. Thus, the computationally intensive part of the LO algorithm (evaluation of a Q function) is avoided while retaining its adaptation property.

VI. Conclusion

We have given a precise definition for the adaptation of handoff algorithms to changing system parameters. We have shown that under our definition of adaptation, the LO algorithm adapts to changing system parameters simply by fixing tradeoff parameter c appropriately. The adaptation rule is explicit and does not require lookup tables. We have also constructed estimators of system parameters based solely on received pilot signal strength samples. Also, information about the mean signal strength and the statistics of shadow fading need not be provided by the base station; rather, they are estimated using the pilot signal strength samples themselves.

We have also presented an easily implementable hysteresis-threshold approximation to the analytically derived LO algorithm, and shown that it retains the adaptive nature of the LO algorithm. Clear rules for selecting the hysteresis and threshold levels as a function of the system parameter estimates have been developed. There is no need to construct lookup tables for adjusting these levels in order to achieve adaptation.

While we have considered only the case of fixed interference levels, both the LO and the hysteresis-threshold approximation algorithms can adapt to changes in the interference levels on the traffic channels by changing Δ. Information about the interference levels may be available at the base station, or may be measured at the mobile.

Finally, it is of great interest to extend the results of this paper to the problem of soft handoff. Research is currently underway to address this extension [12].

Appendix

As seen in Section III, estimation of changing system parameters is crucial to the adaptation of handoff algorithms. Different estimation strategies of varying complexity are discussed here, and a strategy may be selected based on implementation complexity constraints.

The system model under consideration is described by the system parameters $\{v, \overline{d}_i, \sigma_i, \mu_i, \eta_i\colon\ i = 1, 2\}$ (see Section II). For the LO algorithm [see (15)], the set of relevant system parameters is $\mathcal{S} = \{a_i, \sigma_i, \overline{P}_i\colon\ i = 1, 2\}$, where $a_i = \exp(vt_s/\overline{d}_i)$ includes information about v and $\overline{d}_i$, and $\overline{P}_i$ includes information about $\mathrm{E}[X_{k+1,i}|I_k]$ (5). We consider estimators for these system parameters based only on the received pilot strength samples.

A. Estimation of Correlation Coefficient a_i

Consider estimation of the correlation coefficient a_i that is defined in (4). The only observations are the sequences of received signal strengths $\{X_{k,i}\}$, and no knowledge of the local mean pilot powers $\overline{P}_{k,i}$ is assumed. For notational ease, we omit the base station index i in the remainder of this section. Consider a window of M consecutive samples,[3] over which $\overline{P}_k$ can be assumed to be constant. This assumption is valid for small M, such that the local mean pilot strength does not vary much over the time interval Mt_s (distance Mvt_s). Then $\{X_k\}$ will be

[3]For notational ease, we assume that the first sample is at time $k = 1$, although in general the samples may begin at any time.

a (nonzero mean) AR-1 process with correlation coefficient a. Let $\overline{X}$ be the sample mean over the M samples, and let $\overline{\sigma}^2$ be the sample variance. Then the sample correlation $\hat{a}^S$ of these M observations serves as a simple estimator for a, i.e.,

$$\hat{a}^{\mathrm{S}} = \frac{C}{\overline{\sigma}^2} \tag{22}$$

where

$$C = \frac{1}{M-1} \sum_{j=1}^{M-1} (X_j - \overline{X})(X_{j+1} - \overline{X})$$

and

$$\overline{X} = \frac{1}{M} \sum_{j=1}^{M} X_k. \tag{23}$$

Below, we discuss an estimator of shadow fading variance $\hat{\sigma}^{\mathrm{J}}$. Using $\hat{\sigma}^{\mathrm{J}}$ instead of the sample variance in (22) gives the estimator $\hat{a}^{\mathrm{J}} = C/(\hat{\sigma}^J)^2$.

A maximum likelihood (ML) estimator $\hat{a}^{\mathrm{ML}}$ for a can also be found. To derive the ML estimator, assume that the local mean pilot power is constant (equal to $\overline{P}$) and known over the window. Define

$$Y_j = (X_j - \overline{P})/\sigma. \tag{24}$$

Then $\{Y_j\}$ is a zero mean, unit variance AR-1 process. In [8] it is shown that $\hat{a}^{\mathrm{ML}}$ for such a process is the root of the cubic

$$(1+a^2)\left(\sum_{j=1}^{M-1} Y_j Y_{j+1}\right) - a\left(Y_1^2 + 2\sum_{j=2}^{M-1} Y_j^2 + Y_M^2\right) + a(M-1)(1-a^2) = 0. \tag{25}$$

In the practice, however, the local mean power $\overline{P}$ as well as the shadow fading variance σ^2 may not be known, and we may use $Y_j = (X_j - \overline{X})/\hat{\sigma}^{\mathrm{J}}$ in place of (24). In [13], where estimation of AR-1 process parameters is considered in the context of dairy science, a joint ML estimator for σ and a is also shown to result in a cubic for a. For a detailed discussion on the solution of this cubic, and the selection of its appropriate root, see [13, Appendix A].

Fig. 9 shows the probability density function (pdf) of estimators $\hat{a}^{\mathrm{S}}$, $\hat{a}^{\mathrm{ML}}$, and $\hat{a}^{\mathrm{J}}$ for different mobile speeds (different a). To reduce computation (especially in the case of $\hat{a}^{\mathrm{ML}}$), the estimator is computed only once every M samples. As expected, the estimators have a lower variance for small velocities because the local mean pilot strength varies less over smaller distances. The ML estimator outperforms both $\hat{a}^{\mathrm{S}}$ and $\hat{a}^{\mathrm{J}}$. However, evaluating $\hat{a}^{\mathrm{ML}}$ involves solving a cubic, and is therefore computation intensive. We use the estimator $\hat{a}^{\mathrm{J}}$ for the LO algorithm, because it is simpler to evaluate, and provides acceptable performance (Section IV).

B. Estimation of Shadow Fading Variance

As in the previous subsection, consider a block of M samples of pilot signal strength measurements. The sample variance of this block of observations is an estimate of the shadow fading variance $\hat{\sigma}^S$

$$\hat{\sigma}^{\mathrm{S}} = \frac{1}{M-1} \sum_{i=1}^{M} (X_i - \overline{X})^2. \tag{26}$$

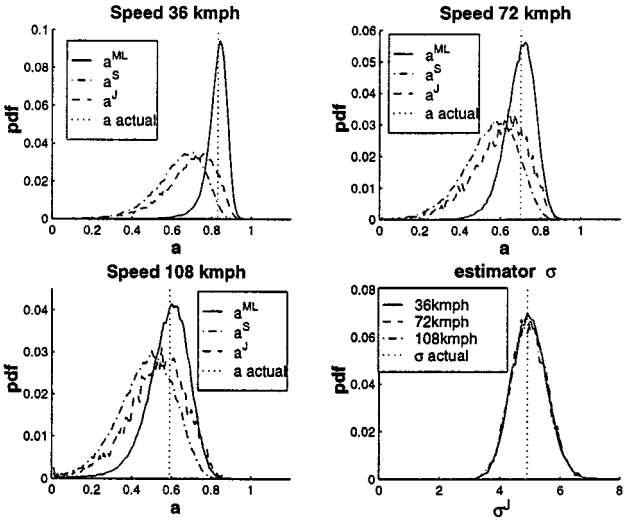

Fig. 9. The probability distribution functions of different estimators of a and σ for various speeds.

When $\{X_k\}$ are independent, the above estimator is unbiased. However, in our case, due to correlation between samples, $\hat{\sigma}^S$ is a biased estimator of σ.

In [5], an estimator for σ was constructed using the quantity R defined below

$$R = \frac{1}{M-1} \sum_{i=1}^{M-1} (X_i - X_{i+1})^2. \tag{27}$$

We evaluate $\mathrm{E}[(X_i - X_{i+1})^2]$ to compute $\mathrm{E}[R]$. Under the assumption that the change in mean signal level, $\overline{P}_k - \overline{P}_{k+1}$ is negligible, it is easy to show that

$$\mathrm{E}[(X_i - X_{i+1})^2] = 2\sigma^2(1-a).$$

This implies from (27) that

$$\mathrm{E}[R] = 2\sigma^2(1-a). \tag{28}$$

Thus, $R/(2(1-a))$ is an unbiased estimator of σ^2. This unbiased estimator can be implemented only when a is known. If we use the estimator $\hat{a} = C/\hat{\sigma}^2$ given in (22) we get

$$(\hat{\sigma}^{\mathrm{J}})^2 = \frac{R}{2\left(1 - \dfrac{C}{(\hat{\sigma}^{\mathrm{J}})^2}\right)}.$$

The above equation, when solved for $\hat{\sigma}^{\mathrm{J}}$, gives

$$\hat{\sigma}^{\mathrm{J}} = \sqrt{C + R/2} \tag{29}$$

with the letter J denoting the joint estimation of σ and a. In Fig. 9 (bottom right), we show the performance of the estimator $\hat{\sigma}^{\mathrm{J}}$. It can be seen that for the given range of mobile speeds, this estimator provides estimates of σ within 1 dB of the actual value.

C. Estimation of Signal Strength

The LO algorithm (15) requires the estimation of $\mathrm{E}[X_{k+1,i}|I_k]$ based on the information I_k. In case the local mean powers $\overline{P}_k$ are known, perfect estimation is possible using (5). However, in a real system, both $\overline{P}_k$ and a are unknown. As above, assume that the mean pilot signal strength does not change much over the observation window. Then, it

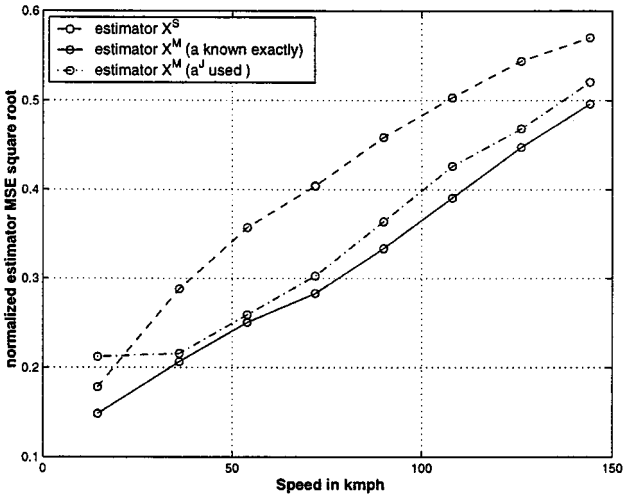

Fig. 10. The MSE of various estimators $\hat{X}_{k+1}$.

is possible to approximate $\overline{P}_k$ by $\overline{P} = \overline{X}$, and a by $\hat{a}^{\mathrm{J}}$. The resulting estimator for X_{k+1} is

$$\hat{X}^{\mathrm{M}}_{k+1} = \overline{P} + \hat{a}^{\mathrm{J}}(X_k - \overline{P}). \tag{30}$$

The simple estimator $\hat{X}^{\mathrm{S}}_{k+1} = X_k$ was used for the numerical results in [2]. This estimator is easier to implement than the one in (30) because it does not require the computation of the mean pilot signal strength.

Fig. 10 shows the MSE of both estimators $\hat{X}^{\mathrm{M}}_{k+1}$ and $\hat{X}^{\mathrm{S}}_{k+1}$ as the mobile speed changes. The MSE is defined as

$$\mathrm{MSE} = \mathrm{E}\left[(\hat{X}_{k+1}(I_k) - \mathrm{E}[X_{k+1}|I_k])^2\right].$$

Normalization is done by dividing the MSE by $\sigma^2(1-a^2)$ [this is the conditional variance of X_{k+1} in (5)]. It is seen that the estimator performance deteriorates as the mobile speed increases. This is because $\overline{P}_k$ changes significantly from sample to sample when mobile speed is higher, leading to a violation of the assumption of constant $\overline{P}_k$ over a block of M samples. It can also be seen that the estimator $\hat{X}^{\mathrm{S}}_{k+1}$ has an MSE that is about 30% greater than that of $\hat{X}^{\mathrm{M}}_{k+1}$. Depending on implementation complexity constraints, either of the estimators $\hat{X}^{\mathrm{S}}_{k+1}$ or $\hat{X}^{\mathrm{M}}_{k+1}$ can be used in a real system.

REFERENCES

[1] G. P. Pollini, "Trends in handover design," *IEEE Commun. Mag.*, vol. 34, pp. 82–90, Mar. 1996.
[2] V. V. Veeravalli and O. E. Kelly, "A locally optimal handoff algorithm for cellular communications," *IEEE Trans. Veh. Technol.*, vol. 46, pp. 603–610, Aug. 1997.
[3] D. Bertsekas, *Dynamic Programming*. Upper Saddle River, NJ: Prentice-Hall, 1987.
[4] R. Vijayan and J. M. Holtzman, "A model for analyzing handoff algorithms," *IEEE Trans. Veh. Technol.*, vol. 42, pp. 351–356, Aug. 1993.
[5] A. Sampath and J. M. Holtzman, "Adaptive handoffs through the estimation of fading parameters," in *Proc. Int. Conf. Commun.*, New Orleans, LA, May 1994, pp. 1131–1135.
[6] M. D. Austin and G. L. Stuber, "Velocity adaptive handoff algorithms for microcellular systems," in *Proc. Int. Conf. Univ. Pers. Commun.*, Ottawa, Canada, Oct. 1993, pp. 793–797.
[7] N. Benvenuto and F. Santucci, "A least squares path-loss estimation approach to handover algorithms," *IEEE Trans. Veh. Technol.*, vol. 48, pp. 437–447, Mar. 1999.
[8] R. Prakash, "Analysis of handoff algorithms," Master's thesis, School of Electrical and Computer Engineering, Cornell Univ., Ithaca, NY, Aug. 1999.
[9] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electron. Lett.*, vol. 27, no. 23, pp. 2145–2146, Nov. 1991.
[10] R. Prakash and V. V. Veeravalli, "Accurate performance analysis of hard handoff algorithms," in *Proc. IEEE Int. Symp. Pers., Indoor Mobile Radio Commun.*, Boston, MA, Sept. 1998, pp. 360–364.
[11] N. Zhang and J. M. Holtzman, "Analysis of handoff algorithms using both absolute and relative measurements," in *Proc. IEEE 44th Veh. Technol. Conf.*, Stockholm, Sweden, June 1994, pp. 82–86.
[12] R. Prakash and V. V. Veeravalli, "Locally optimal soft handoff algorithm," in *Proc. IEEE Veh. Technol. Conf.*, vol. 2, Tokyo, Japan, May 2000, pp. 1450–1454.
[13] K. M. Wade, "Parameter estimation for mixed linear models with first order autoregressive covariance structure," Ph.D. thesis, College Agriculture Life Sciences, Cornell Univ., 1990.

Rajat Prakash (S'97) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Kanpur, in 1997, and the M.S. degree from Cornell University, Ithaca, NY, in 1999. He is currently in the Ph.D. degree program at the University of Illinois at Urbana-Champaign.

His research interests include wireless networks, radio resource management, and mobile communication systems.

Venugopal V. Veeravalli (S'86–M'92–SM'98) received the Ph.D. degree in 1992 from the University of Illinois at Urbana-Champaign, the M.S. degree in 1987 from Carnegie-Mellon University, Pittsburgh, PA, and the B.Tech. degree in 1985 from the Indian Institute of Technology, Bombay, (Silver Medal Honors), all in electrical engineering.

He joined the University of Illinois at Urbana-Champaign in 2000, where he is currently an Associate Professor in the Department of Electrical and Computer Engineering, and a Research Associate Professor in the Coordinated Science Laboratory. He was a Postdoctoral Fellow at Harvard University, Cambridge, MA, during 1992–1993, an Assistant Professor at the City College of NY during 1993–1994, a Visiting Assistant Professor at Rice University, Houston, TX during 1994–1996, and an Assistant Professor at Cornell University, Ithaca, NY during 1996–2000. His research interests include mobile and wireless communications, detection and estimation theory, and information theory.

Dr. Veeravalli is currently an Associate Editor for IEEE JOURNAL OF SELECTED AREAS IN COMMUNICATION—Wireless Communication Series, and an Editor for *Communications in Information and Systems* (CIS). Among the awards he has received for research and teaching are the IEEE Browder J. Thompson Best Paper Award in 1996, the National Science Foundation CAREER Award in 1998, the Presidential Early Career Award for Scientists and Engineers (PECASE) in 1999, and the Michael Tien Excellence in Teaching Award from the College of Engineering, Cornell University, in 1999.