TECHNICAL RESEARCH REPORT

Stochastic Control of Handoffs in Cellular Networks

by R. Rezaiifar, A.M. Makowski, S. Kumar

CSHCN T.R. 94-6
(ISR T.R. 94-73)

STOCHASTIC CONTROL OF HANDOFFS IN CELLULAR NETWORKS

RAMIN REZAIIFAR, ARMAND M. MAKOWSKI * SRIKANTA KUMAR †

Abstract

A Dynamic Programming formulation is used to obtain an optimal strategy for the handoff problem in cellular radio systems. The formulation includes the modeling of the underlying randomness in received signal strengths as well as the movements of the mobile. The cost function is designed such that there is a cost associated with switching and a reward for improving the quality of the call. The optimum decision is characterized by a threshold on the difference between the measured power that the mobile receives from the base stations. Also we study the problem of choosing the "best" fixed threshold that minimizes the cost function. The performance of the optimal and suboptimal strategies are compared.

1 Introduction

Wireless networks are experiencing rapid growth, a trend likely to continue in the foreseeable future. In both micro and macro cellular networks a key issue for efficient operation is the problem of handoffs. A call on a portable/mobile leaving one cell (radio coverage area) and entering a neighboring cell must be transferred to the base station of this neighboring (new) cell. Each handoff involves a signaling cost. Because of the statistical fluctuations in signal strength due to fading, a call may get bounced back and forth between neighboring base stations before it is either successfully handed off or forced to terminate as signal strength falls below acceptable levels. An improperly designed handoff algorithm can result in an unacceptably high level of bouncing (resulting in high signaling costs) and/or a high probability of forced termination. A closely related problem is that of location area updating, wherein a portable/mobile, even though not active, must select and keep reporting to a base station best suited for making and receiving calls. We argue that

*Department of Electrical Engineering and Institute For Systems Research, University of Maryland, College Park, MD 20742. The work of these authors was supported partially through NSF Grant NSFD CDR-88-03012 and through NASA Grant NAGW277S.

†Electrical Engineering and Computer Science, Northwestern University, Evanston, IL 60208.

approaching the handoff problem in a stochastic control framework is most appropriate. We use a Markov decision process formulation, and derive optimal handoff strategies via Dynamic Programming.

Typically, in a cellular mobile communication network (analog or digital), each cell is assigned a separate set of channels (frequencies, carriers, or time slots). The assigned set depends on the frequency planning strategy used for spatial reuse, and maybe fixed or changing dynamically. A successful handoff entails not only the availability of a channel in the new cell (to which the mobile enters) but also an acceptable level of signal strength on the available channel.

To focus mainly on the handoff issue, we take a simple model of two adjacent cells with one channel per cell, and analyze the optimal handoff when a single mobile with an active call moves from one cell to the other. We assume that these channels are always available, distinct, and that their statistical characteristics are independent. Each of these channels are assumed to provide a two-way link between the respective base and the mobile (and thus we do not distinguish between frequency or time division duplex link to achieve this two-way communication). We analyze mobile controlled handoff in the sense that the signal strength on each of these channels is measured periodically at regular intervals at the mobile/portable. The signal strength so measured is subject to both path loss and shadow fading. Handoff decisions are made at these measurement instants. Multipath effects are ignored here as the correlation time is typically much smaller than the measurement interval for most cases of practical interest. Possible interference due to other calls being on a co-channel (e.g., same frequency at another base) are also ignored. Nevertheless, the results derived here form a basis for analyzing enriched models that include such interference, availability of multiple channels, and base station controlled or base-mobile negotiated handoffs.

Our model formulation includes modeling the movement of the mobile as well as the underlying randomness, induced by the (spatially correlated) fading environment, in the signal strengths as observed at the measurement instants.

An optimal handoff strategy should reflect the optimal tradeoff between the call quality (higher signal strength implies a higher a call quality) and the signaling costs. If the handoffs could be accomplished without cost (no signaling costs), the best strategy, trivially, is for the mobile to connect to the base (channel) with higher signal strength at each instant. In the presence of non–zero signaling cost, the best handoff strategy should reflect the optimum intertemporal tradeoff (during the lifetime of the call) between the total signaling costs and the quality or signal strength achievable by the connection, instant to instant, relative to the alternative connection present. Accordingly, for purposes of optimization, we define a cost function that entails a fixed signaling cost for each handoff, and a cost proportional to the power gain foregone when a switch to the higher power is not undertaken. The specific cost function we define, while reflecting the necessary concerns, also simplifies the numerical computations to obtain the threshold. However, the methodology is applicable to other definitions of cost.

We show that the optimal handoff strategy is characterized by a threshold policy. The threshold is defined over the signal strength difference observed on the channels. We then specialize the results to the case of (correlated) log normal fading, a case of practical interest, and compare the performance of the optimum strategy to the best constant threshold policy (hysteresis), often employed in current practice. Conditions for which hysteresis policies do and do not perform well are analyzed.

Much of the previous research on handoffs is based on simulation studies, while the theoretical studies have focused on analyzing the number of handoffs for a given hysteresis strategy [3], [10]. We believe this paper is the first attempt to address handoffs in a control-theoretic framework, and that such an approach will lead to good handoff algorithm design. We have recently become aware of one other study [5] which uses stochastic Dynamic Programming to optimize resources for location area updates; this is, however, a significantly different problem than the handoff issue studied here.

The paper is organized as follows: In Section 2 we present a general Markov decision theoretic framework for addressing the handoff issue. Section 3 introduces the issues related to the movement of mobile and its dynamics. Section 4 discusses the models being used in this work to characterize the stochastic behavior of the received powers as well as some more general models that can be exploited in the same fashion without a significant change in the proposed scheme. The cost function is defined in Section 5 and the corresponding Dynamic Programming formulation is presented in Section 6. The invariance properties of the Dynamic Programming operator are studied in Section 7. These properties allow us to characterize the structure of the optimal policy in Section 8. To asses the effectiveness of the different schemes, we consider the call quality and number of handoffs as two possible measures and Section 9 studies the problem of computing these two quantities once the handoff strategy is set. Finally Section 10 contains several numerical results and comparison between different handoff schemes. Several proofs are relegated to the Appendix. For lack of space, many of the proofs and technical details have been omitted.

A few words on the notation used throughout: For any x in $\mathbb{R}^2$, we write $\|x\|$ for its Euclidean norm. The symbol $\equiv$ stands for defining equality. For any pair of random variables (X, Y), the notation $X =_{st} Y$ means that X and Y have the same distribution. Moreover, $[X \mid Y]$ refers to any random variable which is distributed according to the conditional distribution of X given Y; a similar notation is used for $[X \mid Y = y]$. For any sequence of random variables $\{\xi_t,\ t = 0, 1, \ldots\}$, we set $\xi^t \equiv (\xi_0, \xi_1, \ldots, \xi_t)$ for the history of the sequence up to time $t = 0, 1, \ldots$. The indicator function of any set A is denoted by $\mathbf{1}[A]$.

2 The Model

We now introduce a Markov decision process formulation for the handoff problem faced by a mobile which receives signals from two distinct base stations, labeled base stations zero and one, while moving within a given geographical area. At any given time the mobile has to select the base station through which wireless communication will be achieved. Control information is gathered at sampled epochs and decisions are then taken at these instants. Therefore, under a reasonable assumption of uniform sampling rates, all dynamical processes of interest can be modeled as discrete time processes along the time horizon $t = 0, 1, \ldots$.

2.1 The Underlying Randomness

We begin by describing the elements of the model which are unaffected by the mobile's control actions. This includes randomness in signal propagation and fading as well as possible randomness in mobile's movement. The mobile moves through a region E of the plane $\mathbb{R}^2$, which we assume composed of a finite number of points in the plane. This is done in order to simplify the discussion, with the understanding that most of the developments herein applies to the case of more general regions. The mobile then travels through E according to a stochastic process $\{S_t,\ t = 0, 1, \ldots\}$ with S_t denoting the position in E of the mobile at the beginning of the time slot $[t, t+1)$. At time t, the strength of the received signal from base station i is denoted by P_t^i, $i = 0, 1$; it is measured in dB relative to a fixed transmitter power. For notational convenience, we write $P_t \equiv (P_t^0, P_t^1)$ and $X_t \equiv (S_t, P_t)$. The joint evolution of position and power levels $\{X_t,\ t = 0, 1, \ldots\}$ is modeled as a time-homogeneous Markov process with the following structure: First, we assume that the position process $\{S_t,\ t = 0, 1, \ldots\}$ is by itself a time-homogeneous Markov process on E with one-step transition probability matrix $Q \equiv (Q(s; s'))$ such that

$$\begin{aligned} \mathbf{P}[S_{t+1} = s_{t+1} \mid X^t = x^t] &= \mathbf{P}[S_{t+1} = s_{t+1} \mid S_t = s_t] \\ &= Q(s_t; s_{t+1}). \end{aligned} \tag{2.1}$$

Next, we postulate

$$\begin{aligned} &\mathbf{P}[P_{t+1} \le p \mid X^t = x^t, S_{t+1} = s_{t+1}] \\ &= \mathbf{P}[P_{t+1} \le p \mid X_t = x_t, S_{t+1} = s_{t+1}] \\ &= G(p \mid s_t, p_t, s_{t+1}), \quad p \in \mathbb{R}^2 \end{aligned} \tag{2.2}$$

where $G(\cdot \mid s_t, p_t, s_{t+1})$ denotes the conditional probability distribution of P_{t+1} given that the mobile is in position s_t and s_{t+1} at time t and $t+1$, respectively, and that power strengths at time t were observed at levels p_t. The assumption (2.2) attempts to model the dependence between measured power levels as rather short-term and short-range. Although not entirely accurate, it is nevertheless compatible with modeling assumptions used in previous works [3], [4], [11], [10]; we shall return to this point in Section 4.

Finally, upon combining (2.1) and (2.2) , we see by a simple conditioning argument that

$$\begin{aligned} & \mathbf{P}[S_{t+1} = s_{t+1}, P_{t+1} \le p \mid X^t = x^t] \\ = \quad & \mathbf{E}[\mathbf{1}[S_{t+1} = s_{t+1}]\, \mathbf{P}[P_{t+1} \le p \mid X^t, S_{t+1}] \mid X^t = x^t] \\ = \quad & \mathbf{E}[\mathbf{1}[S_{t+1} = s_{t+1}] G(p \mid S_t, P_t, S_{t+1}) \mid X^t = x^t] \\ = \quad & G(p \mid s_t, p_t, s_{t+1})\, \mathbf{P}[S_{t+1} = s_{t+1} \mid X^t = x^t] \\ = \quad & G(p \mid s_t, p_t, s_{t+1}) Q(s_t; s_{t+1}) \end{aligned} \tag{2.3}$$

and the process $\{X_t,\ t = 0, 1, \ldots\}$ is indeed a time-homogeneous Markov process on $E \times \mathbb{R}^2$.

The call initiated at time $t = 0$ will last a random number T of time slots. We adopt the traditional assumption that the duration of a call is adequately modeled as an exponential random variable. In line with this standard assumption, in our discrete-time setup we assume that the random variable T is geometrically distributed with

$$\mathbf{P}[T = t + 1] = \rho(1 - \rho)^t, \quad t = 0, 1, \ldots \tag{2.4}$$

for some $0 < \rho < 1$. Alternatively, we may interpret ρ as the hangup probability, so that the call can be terminated in every time slot with probability ρ, and this independently of the duration of the ongoing call. The call duration T is assumed independent of the sequence $\{X_t,\ t = 0, 1, \ldots\}$ as well.

2.2 The Controlled System

Fix $t = 0, 1, \ldots$. At the beginning of the time slot $[t, t+1)$, the mobile is in location S_t, the power strengths from the base stations have been measured at levels P_t^0 and P_t^1, and a decision needs to be taken so as to which base station to use for transmission during the time slot $[t, t+1)$. This action is selected on the basis of available information in a way that we now proceed to define: Let U_t denote the $\{0, 1\}$–valued random variable which encodes the decision taken at time t, i.e., if $U_t = i$, $i = 0, 1$, then base station i is being used during the time slot $[t, t + 1)$. For reasons that will become apparent soon, we set $I_t \equiv U_{t-1}$, so that I_t denotes the base station to which the mobile is attached during the time interval $[t - 1, t)$; we also define I_0 as being arbitrary.

The information available to the decision–maker is described by the random variables $\{H_t,\ t = 0, 1, \ldots\}$ which are defined recursively by

$$H_{t+1} \equiv (H_t, U_t, X_{t+1}, I_{t+1}), \quad t = 0, 1, \ldots \tag{2.5}$$

with $H_0 \equiv (X_0, I_0)$. To determine the successive decisions on the basis of this information pattern, we introduce the following notion of a (control) policy: A policy π is a collection of mappings $\{\pi_t,\ t = 0, 1, \ldots\}$ where for each $t = 0, 1, \ldots$, π_t maps the range of H_t into $\{0, 1\}$, with the interpretation that the base station $\pi_t(h_t)$ is used during the time slot $[t, t+1)$ if $H_t = h_t$. The policy π is said to be a Markov stationary if there exists a single mapping $f : E \times \mathbb{R}^2 \times \{0, 1\} \to \{0, 1\}$ such that $\pi_t(h_t) = f(x_t, i_t)$ with x_t determined through $h_t = (h_{t-1}, u_{t-1}, x_t, i_t)$. The class of all control policies is denoted by $\mathcal{P}$.

Fix a pair (x, i) in $E \times \mathbb{R}^2 \times \{0, 1\}$, and $t = 0, 1, \ldots$. For each policy π in $\mathcal{P}$, we associate a probability measure $\mathbf{P}^\pi_{x,i}$ with the following requirements: First, we require

$$\mathbf{P}^\pi_{x,i}[X_0 = x, I_0 = i] = 1. \tag{2.6}$$

Next, we impose

$$\begin{aligned}
&\mathbf{P}^\pi_{x,i}[S_{t+1} = s_{t+1}, P_{t+1} \leq p, I_{t+1} = i_{t+1} \mid H_t, U_t] \\
&= \delta(i_{t+1}, U_t)\, \mathbf{P}^\pi_{x,i}[S_{t+1} = s_{t+1}, P_{t+1} \leq p \mid H_t, U_t] \qquad (2.7)\\
&= \delta(i_{t+1}, U_t) G(p \mid S_t, P_t, s_{t+1}) Q(S_t; s_{t+1}). \qquad (2.8)
\end{aligned}$$

In (2.7) we have used the equality $I_{t+1} = U_t$, while (2.8) expresses the requirement that the underlying randomness of this model be governed by (2.3), and this independently of the policy in use. Finally, we require

$$\mathbf{P}^\pi_{x,i}[U_t = 1 \mid H_t] = \mathbf{1}[\pi_t(H_t) = 1]. \tag{2.9}$$

The model is fully specified if we further assume the random variable T to be independent of the random variables $\{X_t, U_t,\ t = 0, 1, \ldots\}$ under $\mathbf{P}^\pi_{x,i}$, and this for each policy π in $\mathcal{P}$.

Such specifications, and especially (2.8), amount to casting this controlled system as a Markov decision process with "state" process $\{(X_t, I_t),\ t = 0, 1, \ldots\}$. We refer the reader to the monographs [1], [9] for additional material on Markov decision processes.

The model we have introduced is fairly general and flexible enough to cover many situations of practical interest. We briefly review some of the possibilities in the next two sections.

3 The Mobile Dynamics

The prescription of the one-step transition matrix Q defines a directed graph where the points in E act as vertices, and the edges are the pairs of points (s, s') such that $Q(s, s') > 0$. Typically these edges can be mapped into the physical paths over which the mobile's dynamics is restricted, e.g., streets, walking paths and roads. Figures 1 and 2 provide the graphical representations of the two most common situations. In Fig. 1 we displayed a linear motion which corresponds to traveling along a single highway. Figure 2 is a fully two–dimensional situation which arises when the mobile

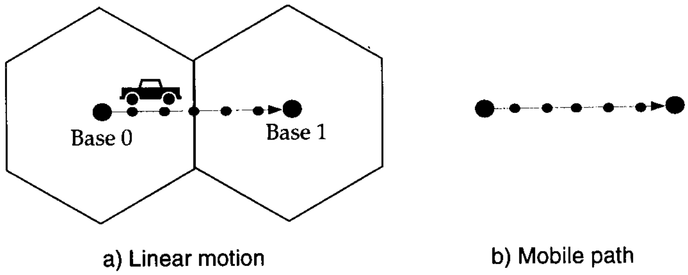

Figure 1: A simple scenario for mobile movement. The dots represent the sampling positions.

user is allowed to move about in a urban area; the edges then represent city blocks or street portions.

This discrete–time model implicitly assumes that a time scale has been postulated, and this in turn determines a (maximal) sampling rate, whereby a sample is collected per time slot. However slower sampling rates can be modeled by simply considering epochs Lt, $t = 0, 1, \ldots$ for some positive integer L, e.g., the sampling rate is now L times slower than the maximal rate. This leads to a model where the original one-step transition matrix Q is replaced by the L–step transition matrix $Q^{(L)}$.

Finally, the assumption that the region E be a discrete subset of $\mathbb{R}^2$ does not constitute an essential restriction for the developments of this paper. Such a constraint does however remove some of technical issues associated with the non–countability of the state space of the Markov decision process. In some situations it might be more appropriate to model E as an arbitrary region of $\mathbb{R}^2$; in that case the one-step transition mechanism is no longer described by the one-step transition matrix Q, but rather by a one-step transition kernel $Q \equiv (Q(s; ds'))$, i.e.,

$$\begin{aligned} \mathbf{P}[S_{t+1} \in B \mid X^t = x^t] &= \mathbf{P}[S_{t+1} \in B \mid S_t = s_t] \\ &= \int_B Q(s_t; ds_{t+1}). \end{aligned} \tag{3.1}$$

This line of inquiry will not be pursued further in this paper due to lack of space.

4 Power Distribution Models

The conditional distribution $G(\cdot \mid s_t, p_t, s_{t+1})$ appearing in (2.2) is the component of the model that is hardest to specify. We devote the present section to the development of a class of models which we shall often consider when carrying out calculations and simulations. These models can be viewed as a dynamic version of a static model which has been widely used to capture shadowing effects [3], [4], [11].

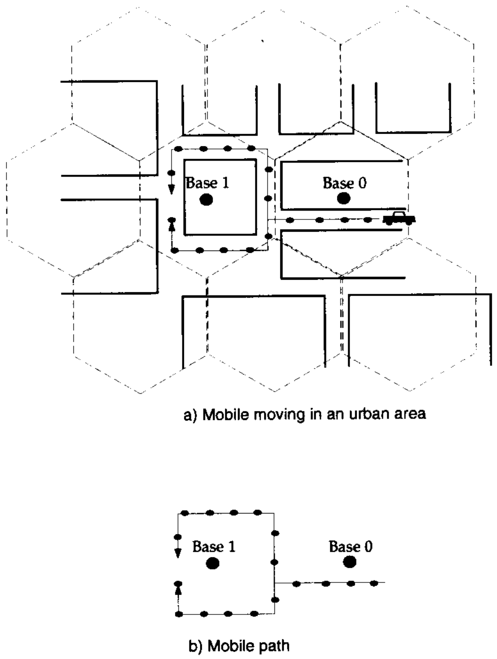

Figure 2: A mobile moving in an urban area. The dots represent the sampling positions.

4.1 A Static Model

We begin with a two–dimensional version of the simple correlation model discussed in [3]: Let $\{W^i(r), r \in \mathbb{R}^2\}$ denote a family of jointly Gaussian random variables with zero mean and variance σ_i^2, $i = 0, 1$, with correlation structure of the form

$$\mathbf{E}[W^i(r)W^i(r')] = \sigma_i^2 \exp(-\beta^{-1}\|r - r'\|), \quad r, r' \in \mathbb{R}^2 \tag{4.1}$$

for constants $\beta > 0$ and $\sigma_i^2 > 0$. The two families $\{W^0(r), r \in \mathbb{R}^2\}$ and $\{W^1(r), r \in \mathbb{R}^2\}$ are assumed independent.

Let b_i denote the location of base station i, $i = 0, 1$. In location s, the strength $P^i(s)$ of the signal produced by the base station i is then given by

$$P^i(s) \equiv A_i - B_i \log(\|s - b_i\|) + W^i(s - b_i), \quad s \in \mathbb{R}^2. \tag{4.2}$$

The constant A_i reflects the transmitter power and is a function of transmission frequency and height of the antennas, while B_i, with typical values in the range of 30-40 dB, models the path loss [4]. We shall find it convenient to write $\mu(s) \equiv (\mu_0(s), \mu_1(s))$ where

$$\mu_i(s) \equiv A_i - B_i \log(\|s - b_i\|), \quad s \in \mathbb{R}^2, \ i = 0, 1. \tag{4.3}$$

4.2 A Simple Dynamic Model

The model (4.1)–(4.2) is a spatial one which specifies the distribution of power levels solely as a function of position, and does not seem to fit naturally into the framework of Section 2. As we seek to develop a dynamic model which is compatible with that spatial model, we first consider the following line of reasoning: We shall assume that the shadowing effects are essentially static, i.e., do not vary much over the duration of a call, and are described by the static random fields $\{P^i(r), r \in \mathbb{R}^2\}$, $i = 0, 1$ – these can be thought as being generated at the beginning of time $t = 0$. It then seems reasonable to argue that the power levels at time t are those given by these static random fields evaluated at the position occupied by the mobile at time t. In other words, the power levels $\{P^i_t,\ t = 0, 1, \ldots\}$ can be obtained by "composing" the static random fields $\{P^i(r), r \in \mathbb{R}^2\}$, $i = 0, 1$ with the mobile's motion, namely

$$P^i_t \equiv P^i(S_t) = \mu_i(S_t) + W^i_t, \quad t = 0, 1, \ldots \tag{4.4}$$

where we have set

$$W^i_t \equiv W^i(S_t - b_i), \quad i = 0, 1,\ t = 0, 1, \ldots \tag{4.5}$$

It is also natural to assume that the random field $\{(P^0(r), P^1(r)), r \in \mathbb{R}^2\}$, or equivalently $\{(W^0(r), W^1(r)), r \in \mathbb{R}^2\}$, is independent of the mobile's trajectory $\{S_t,\ t = 0, 1, \ldots\}$.

Under the foregoing assumptions, we would like to check whether the conditional distribution of P_{t+1} given (X^t, S_{t+1}) indeed satisfies (2.2). The detailed calculations are available in [8] where the following results are established: The random variables P^0_{t+1} and P^1_{t+1} are also conditionally independent given (X^t, S_{t+1}), and for $i = 0, 1$, the random variable P^i_{t+1} is conditionally Gaussian given (X^t, S_{t+1}), with

$$[P^i_{t+1} \mid P^t, S_{t+1}] \sim \mathcal{N}\left(\mu_i(S_{t+1}) + \delta^i_{t+1}, \Delta^i_{t+1}\right). \tag{4.6}$$

The exact expressions for the conditional "mean" δ^i_{t+1} and variance Δ^i_{t+1} are not essential for the discussion, but rather that these quantities depend on the *entire* past history $(P^{i,t}, S^{t+1})$. Therefore, the conditional distribution of P_{t+1} given (X^t, S_{t+1}) depends in general on the entire past (X^t, S_{t+1}), rather than on the most recent history (X_t, S_{t+1}) as required by (2.2). Hence, the suggested model (4.4)–(4.5) does not display the requisite Markov property, and cannot be used wholesale for our purposes as we might have hoped.

Undeterred by this state of affairs, we argue that the probabilistic evolution of power levels typically exhibits only *short*-term memory [4], [11], and that the conditional distribution of P_{t+1} given (X^t, S_{t+1}) can be replaced by that of P_{t+1} given (X_t, S_{t+1}) without incurring any loss of statistical significance. With this in mind, we now set out to compute the latter with the hopes of getting clues as to which kind of models are "compatible" with the static model of [4],[11], while still retaining the desired Markov feature. The calculations are carried out in Appendix A.1.

Proposition 4.1 *Under the foregoing assumptions, the following facts hold:*

1. *The random variables W^0_{t+1} and W^1_{t+1} are conditionally independent given (W_t, S_t, S_{t+1}) where we have set $W_t \equiv (W^0_t, W^1_t)$.*

2. *For $i = 0, 1$, the random variable W^i_{t+1} is conditionally Gaussian given (W_t, S_t, S_{t+1}), with*

$$\begin{aligned}[W^i_{t+1} \mid W_t, S_t, S_{t+1}] &=_{st} [W^i_{t+1} \mid W^i_t, S_t, S_{t+1}] \\ &\sim \mathcal{N}(\gamma^i_{t+1}, \Gamma^i_{t+1}).\end{aligned} \tag{4.7}$$

The conditional mean γ^i_{t+1} and variance Γ^i_{t+1} are given by

$$\gamma^i_{t+1} = W^i_t \exp(-\beta^{-1}\|S_t - S_{t+1}\|) \tag{4.8}$$

and

$$\Gamma^i_{t+1} = \sigma^2_i \left[1 - \exp(-2\beta^{-1}\|S_t - S_{t+1}\|)\right]. \tag{4.9}$$

Using the fact that

$$[P_{t+1} \mid X_t, S_{t+1}] =_{st} \mu(S_{t+1}) + [W_{t+1} \mid W_t, S_t, S_{t+1}] \tag{4.10}$$

we conclude that the random variables P^0_{t+1} and P^1_{t+1} are conditionally independent given (X_t, S_{t+1}), and for $i = 0, 1$, the random variable P^i_{t+1} is conditionally Gaussian given (X_t, S_{t+1}), with

$$[P^i_{t+1} \mid X_t, S_{t+1}] \sim \mathcal{N}\left(\mu_i(S_{t+1}) + \gamma^i_{t+1}, \Gamma^i_{t+1}\right). \tag{4.11}$$

4.3 A General Class of Gaussian Models

Motivated by the discussion of the previous section, we propose the following class of dynamic models for power levels: We posit that power levels have the general form

$$P^i_t \equiv \mu_i(S_t) + W^i_t, \quad i = 0, 1, \ t = 0, 1, \ldots \tag{4.12}$$

where for each $t = 0, 1, \ldots$, the random variables W^0_{t+1} and W^1_{t+1} are conditionally independent given (W^t, S^{t+1}), and for $i = 0, 1$, the random variable W^i_{t+1} is conditionally Gaussian given (W^t, S^{t+1}), i.e.,

$$[W^i_{t+1} \mid W^t, S^{t+1}] \sim \mathcal{N}(\gamma^i_{t+1}, \Gamma^i_{t+1}) \quad t = 0, 1, \ldots \tag{4.13}$$

However, taking the position that temporal variations have short–term memory, we require that the conditional mean γ^i_{t+1} and variance Γ^i_{t+1} depend *only* on the variables W^i_t, S_t and S_{t+1}, and have the general form

$$\gamma^i_{t+1} = g_i(W^i_t; S_t, S_{t+1}) \quad \text{and} \quad \Gamma^i_{t+1} = G_i(S_t; S_{t+1}) \tag{4.14}$$

for a choice of mappings $g_i : \mathbb{R} \times E^2 \to \mathbb{R}$ and $G_i : E^2 \to \mathbb{R}_+$. In fact, taking our cue from (4.8)–(4.9), we shall further constrain g_i and G_i to have the special form

$$g_i(w, s) \equiv wr(\|s - s'\|), \quad w \in \mathbb{R}, \quad s, s' \in E \tag{4.15}$$

and

$$G_i(s, s') \equiv \sigma_i^2 \Gamma(\|s - s'\|), \quad s, s' \in E \tag{4.16}$$

for mappings $r : \mathbb{R}_+ \to \mathbb{R}$ and $\Gamma : \mathbb{R}_+ \to \mathbb{R}_+$. In keeping with the underlying assumptions of our discussion, we require

$$r(0) = 1, \quad \Gamma(0) = \varepsilon \geq 0 \quad \text{and} \quad \Gamma(d) \uparrow 1 \ (d \to \infty). \tag{4.17}$$

The choice $r(d) = \exp(-\beta^{-1} d)$ and $\Gamma(d) = 1 - \varepsilon r(d)^2$ represents a special case suggested by (4.8)–(4.9).

Under these assumptions, the random variables P^0_{t+1} and P^1_{t+1} are conditionally independent given (X^t, S_{t+1}), and for $i = 0, 1$, the random variable P^i_{t+1} are conditionally Gaussian given (X^t, S_{t+1}), i.e.,

$$[P^i_{t+1} \mid X^t, S_{t+1}] \sim \mathcal{N}(\mu_i(S_{t+1}) + \gamma^i_{t+1}, \Gamma^i_{t+1}) \quad i = 0, 1, \ t = 0, 1, \ldots \tag{4.18}$$

The additional assumptions (4.15)–(4.16) imply

$$\begin{aligned} \mu_i(S_{t+1}) + \gamma^i_{t+1} &= \mu_i(S_{t+1}) + W^i_t r(\|S_t - S_{t+1}\|) \\ &= \mu_i(S_{t+1}) + (P^i_t - \mu_i(S_t)) r(\|S_t - S_{t+1}\|). \end{aligned} \tag{4.19}$$

5 Cost Function

In order to formulate the handoff problem as a stochastic optimization problem, we need to define a cost structure which quantifies the cost associated with operating the system under any policy in $\mathcal{P}$. Of course there is no unique way of doing so, and as the optimum handoff policy clearly depends on it, we guide our selection by requesting that the corresponding optimal policy displays "good" properties in terms of implementability.

Here, we first select a cost-per-stage $c : \mathbb{R}^2 \times \{0,1\} \times \{0,1\} \to \mathbb{R}$, and for every initial condition (x, i), we define the total cost function

$$J_\pi(x, i) \equiv \mathbf{E}^\pi_{x,i} \left[\sum_{t=0}^{T-1} c(P_t, I_t, U_t) \right], \quad \pi \in \mathcal{P}. \tag{5.1}$$

The problem of interest is then that of finding a policy $\pi^\star$ in $\mathcal{P}$ such that

$$J_{\pi^\star}(x, i) \leq J_\pi(x, i), \quad (x, i) \in E \times \mathbb{R}^2 \times \{0, 1\} \tag{5.2}$$

for every other policy π in $\mathcal{P}$. Such a policy $\pi^\star$, when it exists, is called the optimal (handoff) policy. We shall shortly present conditions under which the total cost (5.1) is well defined and finite.

To settle on a reasonable cost-per-stage c, we argue as follows: Each time the mobile unit chooses a new base station, a database in the switching center is updated to keep track of the mobile's location. Because frequent and unnecessary switches between base stations can be wasteful of system resources, the cost function must be chosen so as to create a trade off between the two possible decisions, namely switching and not switching. One particular cost-per-stage function with this property associates a cost C with switching from one base station to the other, and penalizes the action of not switching by a cost proportional to the difference in signal strength between the alternative base station and the current one. For example, if the mobile unit is connected to base 0 and the strength of the signal from the other base, namely base 1, is higher by $p^1 - p^0$, then we assign the cost $p^1 - p^0$ for not switching to base station 1. The opportunity cost $p^1 - p^0$ encourages the mobile unit to switch to the better base station, whereas the fixed switching cost C creates a trade off. The corresponding cost-per-stage function, c, is given by

$$c(x,i,u) = \begin{cases} C & \text{if } \; i \neq u \\ (-1)^i(p^1 - p^0) & \text{if } \; i = u, x = (s,(p^0,p^1)) \end{cases} \tag{5.3}$$

and it is used in (5.2) for the remainder of the discussion.

A few remarks are in order before discussing the optimization problem (5.1)-(5.3): First of all, throughout the remainder of this paper, we assume that power level distributions are described by one of the Gaussian models discussed in Section 4.3 under the additional constraints (4.15)–(4.17). For technical reasons that will become apparent shortly, we require that the mapping $r : \mathbb{R}_+ \to \mathbb{R}$ entering (4.15) satisfies the condition

$$0 \leq r(d) \leq 1, \quad d \geq 0. \tag{5.4}$$

The important special case $r(d) = \exp(-\beta^{-1}d)$ does satisfy this condition.

Fix $t = 0, 1, \ldots$. Upon writing

$$Z_t \equiv P_t^1 - P_t^0, \tag{5.5}$$

we readily see from the conditional independence that

$$[Z_{t+1} \mid X^t, S_{t+1}] \sim \mathcal{N}(\theta_{t+1}, \Theta_{t+1}). \tag{5.6}$$

The conditional mean and variance are given by

$$\Theta_{t+1} \equiv (\sigma_1^2 + \sigma_0^2)\Gamma(\|S_t - S_{t+1}\|) \tag{5.7}$$

and

$$\theta_{t+1} \equiv \alpha_t Z_t + \beta_t, \tag{5.8}$$

where we have set

$$\alpha_t \equiv r(\|S_t - S_{t+1}\|) \tag{5.9}$$

and

$$\beta_t \equiv \mu_1(S_{t+1}) - \mu_0(S_{t+1}) - (\mu_1(S_t) - \mu_0(S_t))\,\alpha_t. \tag{5.10}$$

The conditional distribution of the difference Z_{t+1} given the entire history (X^t, S_{t+1}) is thus determined solely by (Z_t, S_t, S_{t+1}), and throughout we denote this conditional distribution by $F(\cdot \mid s_t, z_t, s_{t+1})$.

Next, for any Gaussian random variable ξ, we have

$$\mathbf{E}[| \xi - \mu |] = \sqrt{\frac{2\sigma^2}{\pi}} \tag{5.11}$$

if $\xi =_{st} \mathcal{N}(\mu, \sigma^2)$. Therefore, we find that

$$\begin{aligned}
\mathbf{E}_{x,i}[|Z_{t+1}| \mid X^t, S_{t+1}] &\leq \mathbf{E}_{x,i}[|Z_{t+1} - \theta_{t+1}| \mid X^t, S_{t+1}] + |\theta_{t+1}| \\
&\leq \sqrt{\frac{2\Theta_{t+1}}{\pi}} + |\alpha_t||Z_t| + |\beta_t| \qquad (5.12)\\
&\leq A|Z_t| + B \qquad (5.13)
\end{aligned}$$

where the positive constants are given by

$$A \equiv \sup_{s,s' \in E} | r(\|s - s'\|) | \tag{5.14}$$

and

$$B \equiv \sqrt{\frac{2(\sigma_0^2 + \sigma_1^2)}{\pi}} + 2 \sup_{s \in E} | \mu_1(s) - \mu_0(s) | . \tag{5.15}$$

In deriving the expression for B we have made use of the constraint (5.4), and of the fact that $0 \leq \Gamma(d) \leq 1$ for all $d \geq 0$. Note also that (5.4) implies $0 < A \leq 1$.

It is now a simple matter to conclude by induction that for each (x, i) in $E \times \mathbb{R}^2 \times \{0, 1\}$, we have

$$\mathbf{E}_{x,i}[|Z_t|] \leq \begin{cases} A^t|z| + \frac{1-A^{t+1}}{1-A}B & \text{if} \quad A < 1 \\ |z| + (t+1)B & \text{if} \quad A = 1 \end{cases} \tag{5.16}$$

where $z = p^1 - p^0$ is determined through $(x, i) = ((s, p), i)$. Under the enforced independence assumption on the random variable T, we have

$$\begin{aligned}
|J_\pi(x, i)| &\leq \mathbf{E}^\pi_{x,i}\left[\sum_{t=0}^{\infty} \mathbf{1}[T > t]|c(P_t, I_t, U_t)|\right] \\
&= \sum_{t=0}^{\infty} \mathbf{E}^\pi_{x,i}[\mathbf{1}[T > t]|c(P_t, I_t, U_t)|]
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{t=0}^{\infty} \mathbf{P}^{\pi}_{x,i}[T > t]\, \mathbf{E}^{\pi}_{x,i}[|c(P_t, I_t, U_t)|] \\
&= \sum_{t=0}^{\infty} (1-\rho)^t\, \mathbf{E}^{\pi}_{x,i}[|c(P_t, I_t, U_t)|] \\
&\leq \sum_{t=0}^{\infty} (1-\rho)^t\, (C + \mathbf{E}_{x,i}[|Z_t|])
\end{aligned}
\tag{5.17}
$$

Because $A(1-\rho) < 1$, it is plain from (5.16)–(5.17) that for each policy π in $\mathcal{P}$, the cost function $J_\pi(x,i)$ is well defined and finite. A more careful look at these arguments also shows that

$$
J_\pi(x,i) = \mathbf{E}^{\pi}_{x,i}\left[\sum_{t=0}^{\infty} (1-\rho)^t c(X_t, I_t, U_t)\right]. \tag{5.18}
$$

6 Dynamic Programming Formulation

As an immediate consequence of (5.18) the total cost problem (5.1)-(5.3) can be recast as an infinite horizon *discounted* cost problem with discount factor $1-\rho$. The standard machinery of Dynamic Programming therefore applies and leads to a simple characterization of the optimal policy. In the interest of brevity, we gloss over various technical issues associated with the non–countability of the natural state space $E \times \mathbb{R}^2 \times \{0,1\}$ for this MDP; details are available in [8].

The central object of our analysis is the value function associated with (5.18), and the equation it satisfies: For each (x,i) in $E \times \mathbb{R}^2 \times \{0,1\}$, we define the value function by

$$
V(x,i) \equiv \inf_{\pi \in \mathcal{P}} J_\pi(x,i), \tag{6.1}
$$

and for notational convenience, we set

$$
\widetilde{V}(x,i) \equiv \sum\nolimits_{s' \in E} Q(s;s') \int_{\mathbb{R}^2} V(s',p',i) dG(p' \mid x, s'), \quad x = (s,p). \tag{6.2}
$$

The Dynamic Programming equation for the problem at hand is simply

$$
\begin{aligned}
&V(x,i) \\
&= \min\nolimits_{u=0,1} \left\{ c(x,i,u) + (1-\rho)\widetilde{V}(x,u) \right\}, \quad (x,i) \in S \times \mathbb{R}^2 \times \{0,1\}
\end{aligned}
\tag{6.3}
$$

or equivalently,

$$
\begin{aligned}
&V(x,i) \\
&= \min \left\{ C + (1-\rho)\widetilde{V}(x, i \oplus 1), (-1)^i (p^1 - p^0) + (1-\rho)\widetilde{V}(x,i) \right\}
\end{aligned}
\tag{6.4}
$$

where $\oplus$ denotes modulo two addition. Moreover, the optimal policy $\pi^\star$ is a Markov stationary policy which selects to switch in state (x, i) if and only if

$$C + (1-\rho)\widetilde{V}(x, i \oplus 1) \leq (-1)^i (p^1 - p^0) + (1-\rho)\widetilde{V}(x, i). \tag{6.5}$$

The key observation behind many of the developments in discounted Dynamic Programming lies in the fact that the Dynamic Programming equation (6.3) can be interpreted as the fixed point for a nonlinear operator defined on an appropriate Banach space of functions. To make this more precise, we fix ρ in $(0,1)$ and select a constant K such that

$$0 < K < \frac{\rho}{B(1-\rho)} \tag{6.6}$$

where B is the constant given by (5.15). For any mapping $\varphi : E \times \mathbb{R}^2 \times \{0,1\} \to \mathbb{R}$, we set

$$\|\varphi\| \equiv \sup_{x,i} \frac{| \varphi(x,i) |}{1 + K \mid p^1 - p^0 \mid} \tag{6.7}$$

where the supremum is taken over all pairs (x, i) in $E \times \mathbb{R}^2 \times \{0,1\}$. Let $\mathcal{F}$ denote the collection of all Borel measurable mappings $\varphi : E \times \mathbb{R}^2 \times \{0,1\} \to \mathbb{R}$ for which $\|\varphi\| < \infty$. It is well known that (6.7) defines a norm on $\mathcal{F}$, which makes $\mathcal{F}$ into a Banach space. That $\mathcal{F}$ constitutes indeed the natural function space for our problem should be apparent from the fact that for each policy π in $\mathcal{P}$ the cost function J_π is an element of $\mathcal{F}$, and so is the value function V. These conclusions are easy consequences of the bounds (5.16).

Motivated by the form of the Dynamic Programming equation (6.3), we pose the following definitions: For every mapping φ in $\mathcal{F}$, we associate $\mathbb{R}$–valued mappings $\widetilde{T}\varphi$ and $T_u\varphi$, $u = 0, 1$ defined on $E \times \mathbb{R}^2 \times \{0,1\}$ by setting

$$(\widetilde{T}\varphi)(x,i) \equiv \sum_{s' \in E} Q(s;s') \int_{\mathbb{R}^2} \varphi(s',p',i) dG(p' \mid x, s'), \quad x = (s,p). \tag{6.8}$$

and

$$(T_u\varphi)(x,i) \equiv c(p,i,u) + (1-\rho)(\widetilde{T}\varphi)(x,u) \tag{6.9}$$

for (x,i) in $E \times \mathbb{R}^2 \times \{0,1\}$. These definitions are well posed under the enforced model assumptions as we now show; the proof is given in Appendix A.2.

Proposition 6.1 *Under the model assumptions, the following holds:*

1. *For every mapping φ in $\mathcal{F}$, the function $\widetilde{T}\varphi$ given by (6.8) defines an element of $\mathcal{F}$, with*

$$\|\widetilde{T}\varphi\| \leq (1 + KB)\|\varphi\|. \tag{6.10}$$

2. *For every mapping φ in $\mathcal{F}$, the function $T_u\varphi$ given by (6.9) defines an element of $\mathcal{F}$;*

3. *The operator $T_u : \mathcal{F} \to \mathcal{F}$, $u = 0, 1$, is a strict contraction, i.e.,*

$$\|T_u\varphi - T_u\varphi'\| \leq L\|\varphi - \varphi'\|, \quad \varphi, \varphi' \in \mathcal{F} \tag{6.11}$$

for some constant $0 < L < 1$ *given by*

$$L \equiv (1-\rho)(1+KB) \tag{6.12}$$

Next, we introduce the operator $T : \mathcal{F} \to \mathcal{F}$ by setting

$$(T\varphi)(x,i) \equiv \min_{u=0,1} (T_u\varphi)(x,i), \quad (x,i) \in E \times \mathbb{R}^2 \times \{0,1\} \tag{6.13}$$

for every φ in $\mathcal{F}$. This operator permits a rewriting of the Dynamic Programming equation as $V = TV$, so that V is identified as a fixed point for the operator T. In order to take advantage of this fact, we need several properties of T, which are standard [1], [9], [13], and which are summarized below for easy reference:

Proposition 6.2 *Under the model assumptions, the following holds:*

1. *The operator* $T : \mathcal{F} \to \mathcal{F}$ *is a strict contraction with contraction constant* L *given by (6.12);*

2. *The value function* V *(which is an element of* $\mathcal{F}$*) is the only solution of the fixed point equation*

$$\varphi = T\varphi, \quad \varphi \in \mathcal{F}; \tag{6.14}$$

3. *Moreover, for every element* φ *in* $\mathcal{F}$*, the recursive scheme*

$$\varphi_0 = \varphi, \ \varphi_{k+1} = T\varphi_k, \quad k = 0,1,\ldots \tag{6.15}$$

always converges to the value function V *in the sense that* $\lim_k \|\varphi_k - V\| = 0$*, where* $\lim_k \varphi_k(x,i) = V(x,i)$ *for all* (x,i) *in* $E \times \mathbb{R}^2 \times \{0,1\}$*.*

7 Invariance Properties of the Operator T

Key properties of Dynamic Programming operator T are now discussed. First, an element φ in $\mathcal{F}$ is said to belong to $\mathcal{F}^\star$ if there exists a mapping $\varphi^\star : E \times \mathbb{R} \times \{0,1\} \to \mathbb{R}$, such that

$$\varphi(x,i) = \varphi^\star(s, p^1 - p^0, i), \quad (x,i) \in E \times \mathbb{R}^2 \times \{0,1\} \tag{7.1}$$

with the understanding that $(x,i) = (s,p,i)$.

Lemma 7.1 *For every element* φ *of* $\mathcal{F}^\star$*,* $\widetilde{T}\varphi$*,* $T_u\varphi$*,* $u = 0,1$*, and* φ *are all elements of* $\mathcal{F}^\star$*.*

Proof: If φ is an element of $\mathcal{F}^\star$, then $\widetilde{T}\varphi$ is also an element of $\mathcal{F}^\star$ by virtue of (5.7)–(5.10). Because the cost–per–stage c clearly belongs to $\mathcal{F}^\star$, so do $T_u\varphi$, $u = 0,1$, and so ultimately does $T\varphi$. ■

The mappings φ and $\varphi^\star$ entering the definition (7.1) are in many–to–one correspondence with each other. It is therefore natural to adopt the convention that

$\mathcal{F}^\star$ also denotes the class of Borel functions $\varphi : E \times \mathbb{R} \times \{0,1\} \to \mathbb{R}$ such that the mapping $(x,i) \to \varphi(s, p^1 - p^0, i)$ is an element of $\mathcal{F}$. The definition (6.7) reduces to

$$\|\varphi\| \equiv \sup_{s,z,i} \frac{|\varphi(s,z,i)|}{1+K|z|}, \quad \varphi \in \mathcal{F}^\star \tag{7.2}$$

and also yields a norm on $\mathcal{F}^\star$, which turns it into a Banach space.

With this interpretation, the operators $\widetilde{T}$, T_u, $u = 0,1$, and T can now be viewed as acting on $\mathcal{F}^\star$, rather than on $\mathcal{F}$, provided some obvious changes are made: For every φ in $\mathcal{F}^\star$ and every (s,z,i) in $E \times \mathbb{R} \times \{0,1\}$, we set

$$(\widetilde{T}\varphi)(s,z,i) \equiv \sum_{s' \in E} Q(s;s') \int_{\mathbb{R}^2} \varphi(s',z',i) dF(z' \mid s,z,s'), \tag{7.3}$$

$$(T_u\varphi)(s,z,i) \equiv c(z,i,u) + (1-\rho)(\widetilde{T}\varphi)(s,z,u) \tag{7.4}$$

and

$$(T\varphi)(s,z,i) \equiv \min_{u=0,1} (T_u\varphi)(s,z,i). \tag{7.5}$$

Both Propositions 6.1 and 6.2 hold true in this modified context.

Next, we say that a mapping $\varphi : E \times \mathbb{R} \times \{0,1\} \to \mathbb{R}$ is an element of $\mathcal{C}$ if it belongs to $\mathcal{F}$ and if for each s in E and $i = 0,1$, the mapping $z \to \varphi(s,z,i)$ is continuous on $\mathbb{R}$. The key fact of interest here is that $\mathcal{C}$ is invariant under T; a proof is available in [8]. In fact, a little more can be said:

Lemma 7.2 *For every element φ of $\mathcal{C}$, $\widetilde{T}\varphi$, $T_u\varphi$, $u = 0,1$, and $T\varphi$ are all elements of $\mathcal{C}$.*

Finally we conclude with a property which proves crucial in establishing the structure of the optimal policy. For every element φ of $\mathcal{F}^\star$, we set

$$(\Delta\varphi)(s,z) \equiv \varphi(s,z,1) - \varphi(s,z,0), \quad s \in E, z \in \mathbb{R}. \tag{7.6}$$

The element φ of $\mathcal{F}^\star$ is said to belong to $\mathcal{D}$ if for each s in E, the mapping $z \to (\Delta\varphi)(s,z)$ is non–increasing on $\mathbb{R}$. The proof of the next lemma is given in Appendix A.3.

Lemma 7.3 *If φ is an element of $\mathcal{C} \cap \mathcal{D}$, then so are $\widetilde{T}\varphi$ and $T\varphi$.*

8 On the Structure of the Optimal Policy

In this section we develop various results which provide insights into the structure of the optimal policy.

Lemma 8.1 *The value function V is an element of $\mathcal{F}^\star$.*

Proof: We consider the recursive scheme (6.15) with zero initial condition, i.e., $\varphi_0 = 0$. Because the initial condition $\varphi_0 = 0$ is in $\mathcal{F}^\star$, the iterates $\{\varphi_k,\ k = 0, 1, \ldots\}$ are all in $\mathcal{F}^\star$ by Lemma 7.1, and so does V by virtue of Claim **3** of Proposition 6.2. ■

The next result addresses the smoothness of the value function, a fact we shall need for technical reasons later in this section.

Lemma 8.2 *The value function V belongs to $\mathcal{C}$.*

Proof: Again we consider the recursive scheme (6.15) with zero initial condition, i.e., $\varphi_0 = 0$. Because the initial condition $\varphi_0 = 0$ is an element of $\mathcal{C}$, the iterates $\{\varphi_k,\ k = 0, 1, \ldots\}$ are all in $\mathcal{C}$ by Lemma 7.2. On the other hand, by virtue of Claim **3** of Proposition 6.2 we readily see for each $R > 0$ that

$$\lim_k \sup_{|z| \leq R} \frac{|\varphi_k(s,z,i) - V(s,z,i)|}{1 + K|z|} = 0, \quad s \in E,\ i = 0, 1 \tag{8.1}$$

and the convergence $\lim_k \varphi_k(s,z,i) = V(s,z,i)$ is uniform (in z) on compact sets. The continuity of the iterates $\{\varphi_k,\ k = 0, 1, \ldots\}$ now implies that of V. ■

Lemma 8.3 *The value function V and $\widetilde{T}V$ are elements of $\mathcal{C} \cap \mathcal{D}$.*

Proof: We have already shown in Lemma 8.2 that V belongs to $\mathcal{C}$, hence $\widetilde{T}V$ also belongs to $\mathcal{C}$ by Lemma 7.2. Next we again consider the recursive scheme (6.15) with zero initial condition, i.e., $\varphi_0 = 0$. Because the initial condition $\varphi_0 = 0$ is an element of $\mathcal{C} \cap \mathcal{D}$, we conclude by Lemma 7.3 that the iterates $\{\varphi_k,\ k = 0, 1, \ldots\}$ and $\{\widetilde{T}\varphi_k,\ k = 0, 1, \ldots\}$ are all in $\mathcal{C} \cap \mathcal{D}$.

By virtue of Claim **3** of Proposition 6.2 we have $\lim_k (\Delta\varphi_k)(s,z) = (\Delta V)(s,z)$ for all (s,z) in $E \times \mathbb{R}$, and V inherits membership in $\mathcal{C} \cap \mathcal{D}$ from the iterates $\{\varphi_k,\ k = 0, 1, \ldots\}$. A similar argument holds for $\widetilde{T}V$ as it can be shown [8] that $\lim_k (\Delta\widetilde{\varphi}_k)(s,z) = (\Delta\widetilde{V})(s,z)$ for all (s,z) in $E \times \mathbb{R}$; details are omitted in the interest of brevity. ■

We are ready to discuss the structure of the optimal policy. A handoff policy π is said to be a threshold policy with threshold functions $\tau_i : E \to \mathbb{R}$, $i = 0, 1$, if it is a Markov stationary policy such that

$$\pi^\star(s,z,0) = 1 \quad \text{iff} \quad z \geq \tau_0(s), \tag{8.2}$$

and

$$\pi^\star(s,z,1) = 0 \quad \text{iff} \quad z \leq \tau_1(s) \tag{8.3}$$

for every (s,z) in $E \times \mathbb{R}$.

Proposition 8.1 *Under the model assumptions, the optimal handoff policy $\pi^\star$ is a threshold policy with threshold functions $\tau_i^\star : E \to \mathbb{R}$, $i = 0, 1$, which are uniquely determined through the equations*

$$C + (-1)^i(1 - \rho)(\Delta V)(s,z) = (-1)^i z, \quad s \in E,\ i = 0, 1. \tag{8.4}$$

In fact, $\tau_1^\star(s) < \tau_0^\star(s)$ for all s in E.

Proof: Fix (s,z,i) in $E \times \mathbb{R} \times \{0,1\}$. We begin by rewriting the Dynamic Programming equation (6.4) in the form

$$\begin{aligned} V(s,z,i) \\ &= \min\left\{C + (1-\rho)(\tilde{T}V)(s,z,i\oplus 1), (-1)^i z + (1-\rho)(\tilde{T}V)(s,z,i)\right\}. \end{aligned} \tag{8.5}$$

The optimal policy $\pi^\star$ is the Markov stationary policy which selects to switch in state (s,z,i) if and only if

$$C + (1-\rho)(\tilde{T}V)(s,z,i\oplus 1) \leq (-1)^i z + (1-\rho)(\tilde{T}V)(s,z,i), \tag{8.6}$$

or equivalently, if and only if

$$C + (1-\rho)(-1)^i(\Delta\tilde{T}V)(s,z) \leq (-1)^i z. \tag{8.7}$$

By Lemma 8.3, $z \to \Delta\tilde{T}V(s,z)$ is monotone non–increasing and continuous. Hence, for $i = 0$ (resp. $i = 1$) the left hand side of the inequality (8.7) is continuous and monotone non–increasing (resp. non–decreasing), while its right hand side is continuous and *strictly* increasing (resp. decreasing). It is now a simple matter to conclude that the switching sets $B_i(s) \equiv \{z \in \mathbb{R} : C + (1-\rho)(-1)^i(\Delta\tilde{T}V)(s,z) \leq (-1)^i z\}$, $i = 0,1$ are non–empty closed and connected sets which are disjoint (owing to the condition $C > 0$). In fact, $B_0(s) = [\tau_0^\star(s), \infty)$ with $\tau_0^\star(s) = \inf B_0(s)$, and $B_1(s) = (-\infty, \tau_1^\star(s)]$ with $\tau_1^\star(s) = \sup B_1(s)$, and the optimal policy is of threshold type. Because $B_0(s)$ and $B_1(s)$ are disjoint sets, we see that $\tau_1^\star(s) < \tau_0^\star(s)$, and the defining equalities (8.4) are simple consequences of (8.7) and of continuity properties mentioned earlier. ■

For the special case

$$r(d) = 0, \quad d \geq 0 \tag{8.8}$$

some additional properties can be derived for the optimal threshold functions $\tau_i^\star$, $i = 0,1$. When $r(d) = \exp(-\beta^{-1}d)$, this corresponds to $\beta = 0$. We first show that the optimal thresholds in a given position are related to each other in a simple manner.

Corollary 8.1 *Under the condition (8.8), the optimal threshold functions* $\tau_i^\star$, $i = 0,1$, *satisfy the relation*

$$\tau_0^\star(s) = 2C + \tau_1^\star(s), \quad s \in E. \tag{8.9}$$

Proof: Under condition (8.8), the conditional distribution $F(\cdot|s,z,s')$ does not depend on z, hence $(\tilde{T}V)(s,z,i)$ is independent of z for each s in E and $i = 0,1$. Therefore, with a slight abuse of notation, the defining equalities (8.4) reduce to

$$\tau_i^\star(s) = (-1)^i C + (1-\rho)(\Delta\tilde{T}V)(s), \quad i = 0,1, \tag{8.10}$$

and the conclusion follows by direct inspection. ■

Next, we consider the case when the mobile travels along a straight line connecting the two base stations, and this in a unidirectional manner: For sake of concreteness we take $E = \{(0,0), (1,0), \ldots, (N,0)\}$ for some integer N, and assume that for all $j = 0, 1, \ldots, N$, $Q((0,k);(0,j)) = \mathbf{1}[j = k+1]$ for $k = 0, 1, \ldots, N-1$ and $Q((0,N);(0,j)) = \mathbf{1}[j = N]$. We refer the reader to [8] for a proof of the monotonicity of the optimal thresholds:

Corollary 8.2 *Under the condition (8.8), the optimal threshold functions $\tau_i^\star : E \to \mathbb{R}$, $i = 0, 1$, are each monotone non-increasing, i.e., $\tau_i^\star((k+1) \leq \tau_i^\star(k)$ for all $k = 0, 1, \ldots, N-1$.*

9 Average Quality of Call and Expected Number of Handoffs

Once the a handoff policy (be it optimal or not) has been selected, it is of interest to compute the expected value of the quality of the call and the expected number of handoffs that the mobile experiences while the optimal policy is in effect. These two quantities constitute good measures of the effectiveness of a handoff policy. Other criteria include the expected delay in handoff which has been studied by Vijayan and Holtzman [10].

We define the call quality function C_π of the policy π to be the the mean value of the strength of the received signal form the active base stations under the policy π during the call session, namely

$$Q_\pi(x,i) \equiv \mathbf{E}^\pi_{x,i}\left[\sum_{t=0}^{T-1} I_t P_t^1 + (1-I_t)P_t^0\right], \quad (x,i) \in E \times \mathbb{R}^2 \times \{0,1\}. \tag{9.1}$$

On the other hand, the expected number of handoffs under the policy π is defined by

$$S_\pi(x,i) \equiv \mathbf{E}^\pi_{x,i}\left[\sum_{t=0}^{T-1} \mathbf{1}[U_t \neq I_t]\right] \quad (x,i) \in E \times \mathbb{R}^2 \times \{0,1\}. \tag{9.2}$$

An argument similar to that leading to (5.18) yields the alternate expressions

$$C_\pi(x,i) = \mathbf{E}^\pi_{x,i}\left[\sum_{t=0}^{\infty}(1-\rho)^t (I_t P_t^1 + (1-I_t)P_t^0)\right] \tag{9.3}$$

and

$$S_\pi(x,i) = \mathbf{E}^\pi_{x,i}\left[\sum_{t=0}^{\infty}(1-\rho)^t \mathbf{1}[U_t \neq I_t]\right] \tag{9.4}$$

so that both C_π and S_π can be written as discounted cost functions.

For any Markov stationary policy π, and in particular for any threshold policy, this fact can be exploited for numerical purposes by interpreting C_π and S_π as fixed points for suitably defined contractions. More precisely, to evaluate the expected quality of call, for each Markov stationary policy π, we consider an operator $K_\pi : \mathcal{F} \to \mathcal{F}$ of the form

$$(K_\pi \varphi)(x,i) \equiv (K_{\pi(x,i)}\varphi)(x,i), \quad \varphi \in \mathcal{F} \tag{9.5}$$

for every (x,i) in $E \times \mathbb{R}^2 \times \{0,1\}$, where for each $u = 0,1$, the operator $K_u : \mathcal{F} \to \mathcal{F}$ is defined by

$$(K_u\varphi)(x,i) \equiv p^i + (1-\rho)(\widetilde{T}\varphi)(x,u), \quad \varphi \in \mathcal{F}. \tag{9.6}$$

As in Proposition 6.1, the operators, K_u, $u = 0,1$, are contractions on $\mathcal{F}$, and so is K_π. It follows from the Markov property that the call quality function C_π is the unique fixed point of K_π, and can be evaluated through the recursion

$$\varphi_0 = 0, \ \varphi_{k+1} = K_\pi \varphi_k, \quad k = 0,1,\ldots \tag{9.7}$$

by invoking the appropriate version of Claim **3** of Proposition 6.2. To compute the expected number of switches, we use instead the operator $K^\star_\pi : \mathcal{F} \to \mathcal{F}$ which is of the form

$$(K^\star_\pi \varphi)(x,i) \equiv (K^\star_{\pi(s,z,i)}\varphi)(s,z,i), \quad \varphi \in \mathcal{F}^\star \tag{9.8}$$

for every (s,z,i) in $E \times \mathbb{R}^2 \times \{0,1\}$, where for each $u = 0,1$, the operator $K^\star_u : \mathcal{F}^\star \to \mathcal{F}^\star$ is defined by

$$(K^\star_u\varphi)(s,z,i) \equiv \mathbf{1}[u \neq i] + (1-\rho)(\widetilde{T}\varphi)(s,z,u), \quad \varphi \in \mathcal{F}^\star. \tag{9.9}$$

This time, the operators $K^\star_u$, $u = 0,1$, are contractions on $\mathcal{F}^\star$, and so is $K^\star_\pi$. The unique fixed point of $K^\star_\pi$ is S_π, and is obtained through the recursion

$$\varphi_0 = 0, \ \varphi_{k+1} = K^\star_\pi \varphi_k, \quad k = 0,1,\ldots \tag{9.10}$$

by invoking the appropriate version of Claim **3** of Proposition 6.2.

We close this section with the behavior of the optimum cost as a function of the switching cost C. Let $V(x,i,C)$ stands for the value function $V(x,i)$ defined by (6.1) when the switching cost has value C:

Proposition 9.1 *For each (x,i) in $E \times \mathbb{R}^2 \times \{0,1\}$, the value function $V(x,i,C)$ is a concave and nondecreasing function of the switching cost C.*

Proof: Fix (x,i) in $E \times \mathbb{R}^2 \times \{0,1\}$, and let π be an arbitrary policy in $\mathcal{P}$. If $J_\pi(x,i,C)$ denotes the total cost when the switching cost has value C, then direct inspection of (5.18) shows that $J_\pi(x,i,C) = A_\pi(x,i) + CS_\pi(x,i)$, with $S_\pi(x,i)$ given by (9.4) and

$$A_\pi(x,i) \equiv \mathbf{E}^\pi_{x,i}\left[\sum_{t=0}^{\infty}(1-\rho)^t \mathbf{1}[U_t = I_t](-1)^{U_t}(P^1_t - P^0_t)\right]. \tag{9.11}$$

Table 1: NOMINAL PARAMETERS USED FOR NUMERICAL RESULTS

B	σ	ρ	c	β	Distance between bases
30	5 dB	0.2	6	200 m	2Km

Therefore, the mapping $C \to J_\pi(x,i,C)$ is affine and monotone nondecreasing (since $B_\pi(x,i) \geq 0$). The announced conclusion is now immediate from the fact $V(x,i,C) = \inf_{\pi \in \mathcal{P}} J_\pi(x,i,C)$, as we recall that the infimum of affine (resp. monotone non–decreasing) functions yields a concave (resp. monotone non–decreasing) function. ■

10 Numerical Results

In this section we exploit the structure of the proposed handoff strategy in order to obtain the optimum solution for a couple of scenarios which are described below. The discussion is carried out for the special case

$$r(d) = \exp(-\beta^{-1}d) \tag{10.1}$$

and

$$\Gamma(d) = 1 - r(d)^2, \tag{10.2}$$

for all $d \geq 0$ with $\sigma^2 = \sigma_0^2 = \sigma_1^2$.

Scenario 1. We first look at the simple case where the mobile travels on a line connecting the two bases. In Section 8, when $\beta = 0$, we pointed out that the thresholds are non-increasing functions of the position; Figure 3 confirms this fact. The displayed monotonicity of the thresholds corroborates the intuitive belief that were the current base station be base 0, the threshold should be lowered as the mobile gets closer to base station 1, in order to make it easier to switch from base 0 to base 1.

Although information about the distance of the mobile from the base stations is usually not too difficult to obtain [11], it is also possible to find the best fixed threshold (which does not vary with distance). This can be found with the help of a numerical optimization algorithm which seeks the minimum of the cost viewed as a function of fixed thresholds (thus defined on $\mathbb{R}^2$). Figure 4 depicts the cost versus the fixed thresholds H_1 and H_0. The flat surface at the bottom of this figure is the optimum cost. We used a simple steepest descent method to find the minimum of the function. The sub-optimal thresholds are shown in Fig. 3 together with the optimal ones for comparison purposes. The parameters used for this numerical results are given in Table 1.

Scenario 2. Next we consider a more realistic situation of a mobile traveling in a two–dimensional plane as shown in Fig. 5. We also add the possibility that at some point the road divides into two different paths, with the traffic pattern

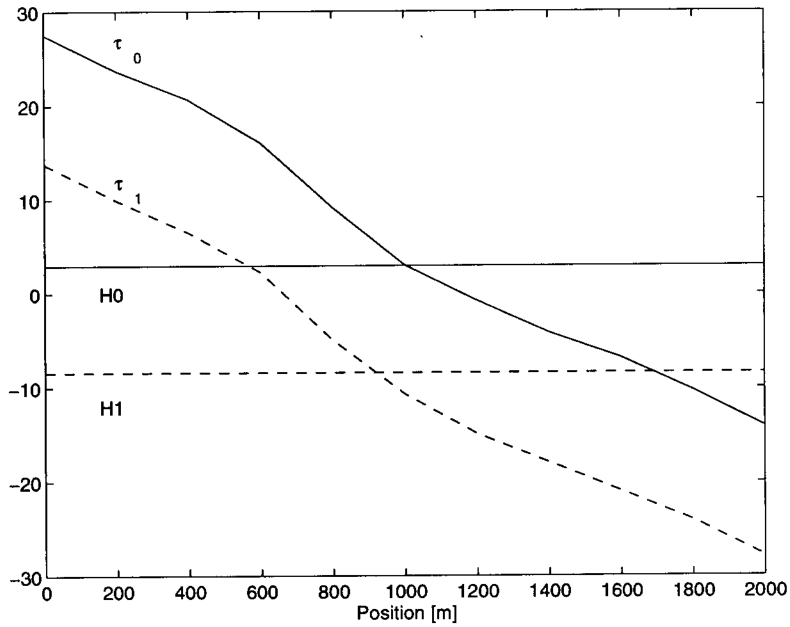

Figure 3: The optimal and sub-optimal thresholds (linear 1-dim motion)

being such that %70 of the mobiles take one path and the rest take the other one. In this case a three–dimensional figure helps to present the optimal thresholds for each location. For each sampling position there are two thresholds $\tau_i, \quad i = 0, 1$. The mobile path and the thresholds are shown in Fig. 6. As was the case for the simple one–dimensional mobile movement, in this scenario the thresholds are lower for the points that are closer to base 1. The optimal and sub-optimal thresholds for the two–dimensional motion are shown in Fig. 7 which is basically another way of presenting the thresholds depicted in Fig. 6. Here however the jump in the threshold function might be misleading. The occurrence of the jump results from the fork-shape of the mobile path, and the jump which occurs at location 12 reflects the fact that there is a significant difference in the distances of location 12 and 13 to base 1 (see Fig. 5).

Clearly, the solution of the optimization problem described here does depend on the structure of the cost function itself, as well as on the choice of the various parameters that enter the cost function. One of the important parameters is the switching cost C, and in what follows we present two methods to pick a reasonable value for this parameter. Note that in the cost function presented in (5.3) the switching cost is being compared with the improvement in the signal strength in dB. We must therefore decide how expensive is the switching action in terms of the amount of improvement that can be achieved by switching to the better base station. Alternatively, the call quality can be computed for different values of C and

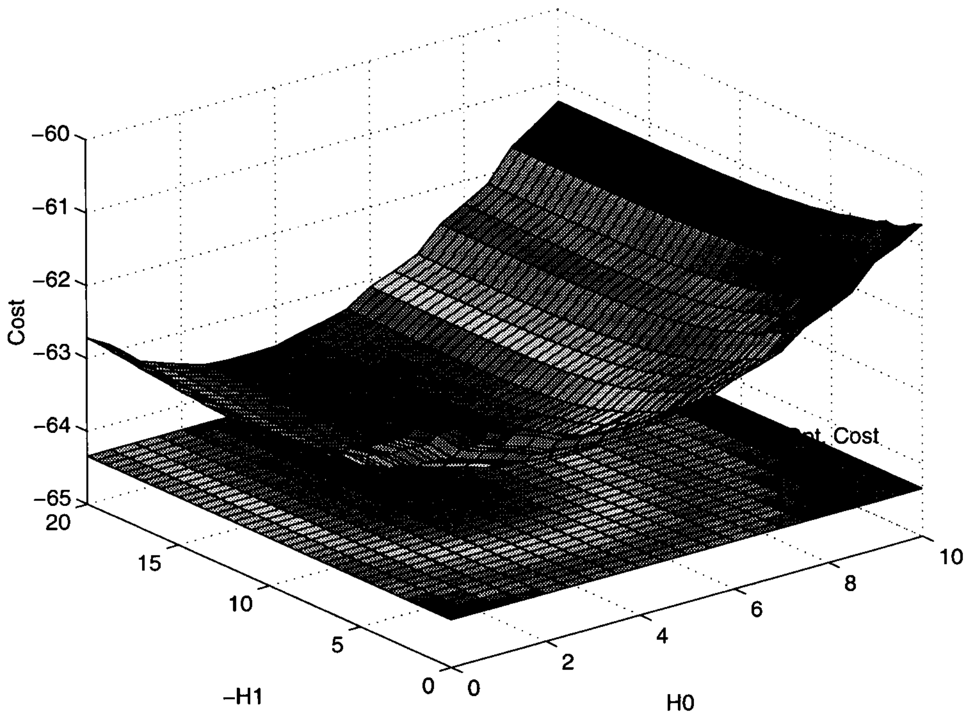

Figure 4: Cost function versus the value of fixed thresholds (linear 1-dim motion).

based on the desired value of the average call quality, the appropriate switching cost can be obtained. In Fig. 8 we have displayed the call quality versus switching cost; for the purpose of normalization, we set $A_i = 0$ so that the constant A_i must be added to the numerical values for the average call quality to obtain the true value. As expected, the call quality degrades as the switching becomes more expensive because it makes the switching action more sluggish. Figure 9 illustrates the effect of changing the various parameters in the problem on the optimal thresholds. It reveals that the optimal solution is very insensitive to the value of the variance σ^2 or of the hangup rate ρ, whereas it is quite sensitive to the correlation factor β.

Finally, we assess the effectiveness of the proposed method by comparing different aspects of three handoff strategies, namely, the optimal policy, the best fixed (sub-optimal) threshold policy, and a non-optimal threshold policy with thresholds equal to the value of σ. The results in Table 2 show that the optimal strategy achieves a better call quality while making fewer switches than the other two strategies. Even the suboptimal strategy shows an improvement over the non-optimal method in both call quality and expected number of switches. In interpreting the expected number of switches we note that $(1-\rho)$ acts as a discount factor, so those switches being made at a later time have less weight than those which occur closer to $t = 0$. It is also worth emphasizing that the optimization scheme creates a balance

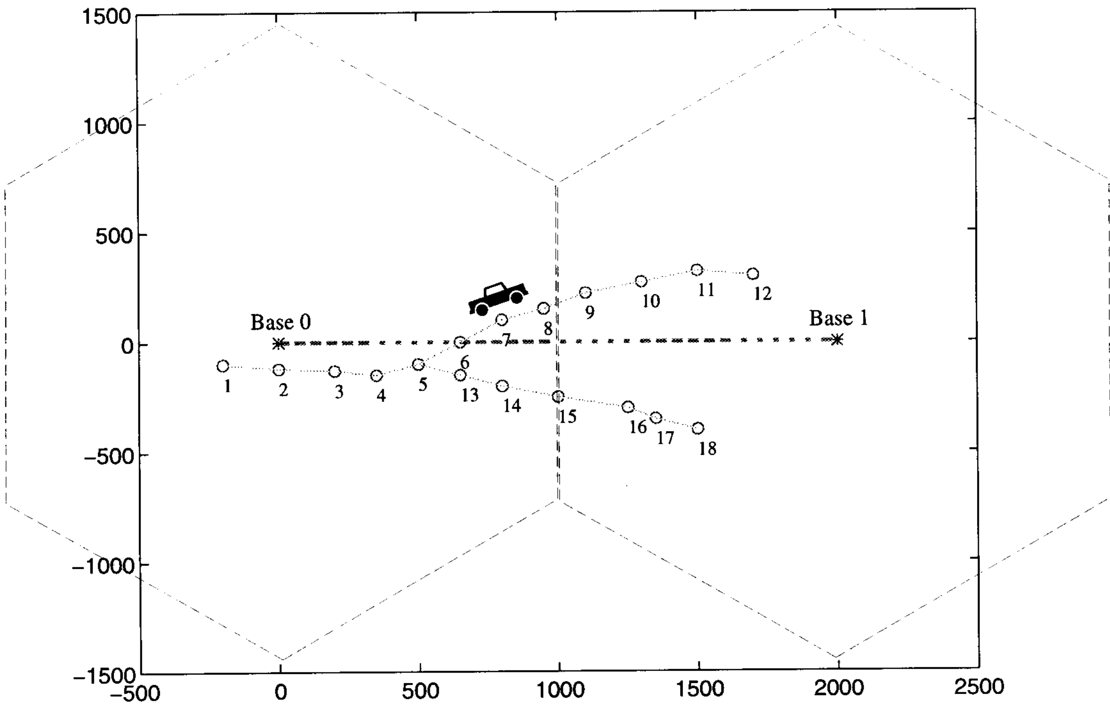

Figure 5: A fork-shape path. At the position 5 mobile chooses one of the paths with a pre-specified probability.

between call quality and number of switches; otherwise we could improve call quality by choosing a very small threshold which obviously has the effect of increasing the number of switches.

11 Conclusions

The problem of handoff in a cellular environment has been cast as a Markov decision problem. We then exploit the well-developed machinery of Dynamic Programming to derive the structure of the optimal handoff policy, and this under an interesting range of model assumptions. The optimal policy is obtained by minimizing a cost function that creates a balance between two conflicting measures, i.e., the number of switches between cell sites, and quality of the call.

The optimal strategy is shown to be of the threshold type, a fact which greatly facilitates its implementation. Through numerical computation we demonstrated that the optimal policy outperforms the conventional non-optimal handoff policy

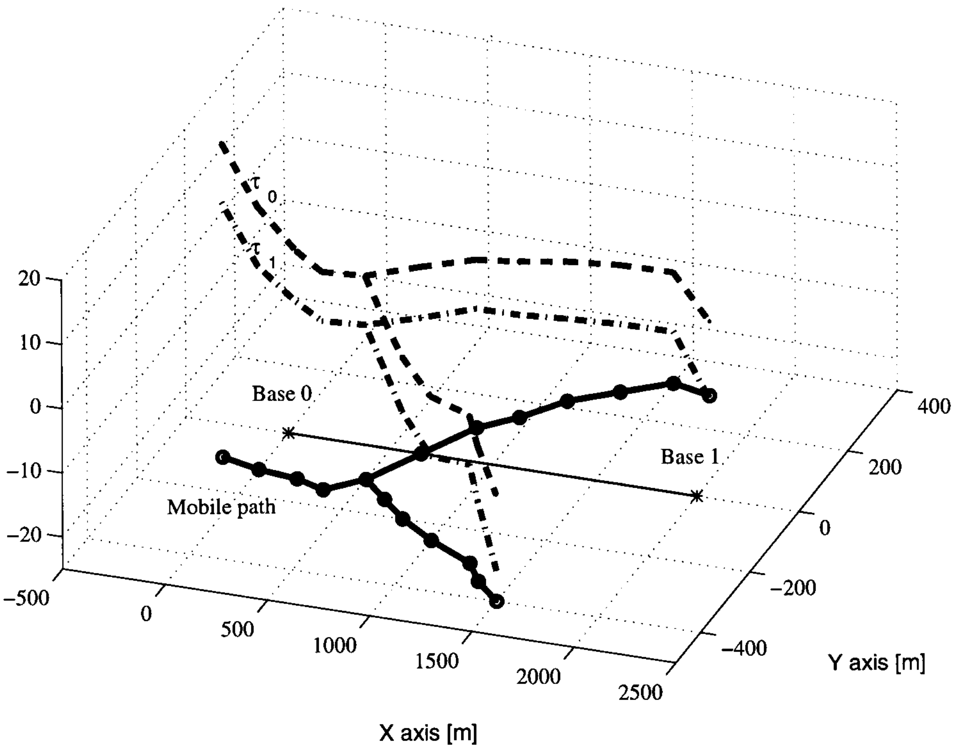

Figure 6: Mobile path together with the optimum thresholds.

in both the number of switches between the cell sites and the quality of the call. The proposed design methodology for handoff policies is also applicable for indoor wireless communication as well as for personal communication systems (PCS); in these situations the size of the cells are much smaller (microcells and picocells) and the use of a sensible handoff policy is even more crucial.

Several extensions of the model studied here will prove useful. The optimal handoff strategy depends on the mobility model. In practice, different mobiles/portables may have different patterns of movement, thus requiring different mobility models, whereas a common handoff strategy may be desired for the system. This aggregation

Table 2: Call Quality, Value of Cost Function, and No. of Switches

	Value of cost function	Avg. call quality	Expected no. of switches
Variable threshold	-19.01	-442.80	0.28
Best fixed threshold	-18.58	-443.60	0.34
σ-threshold policy	-17.49	-444.15	0.63

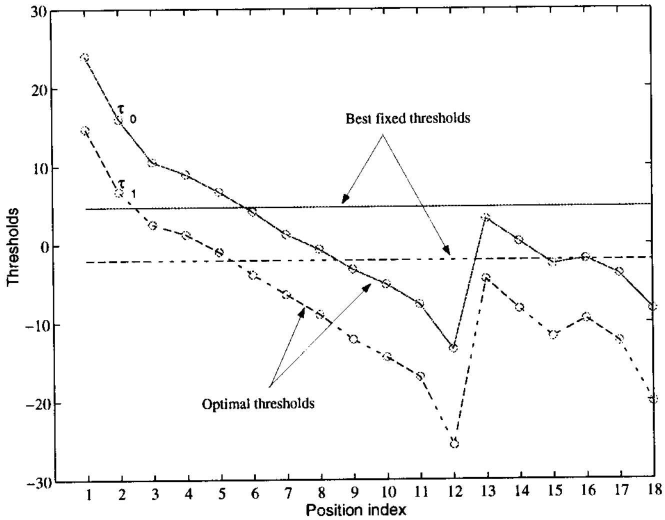

Figure 7: Optimum thresholds for each location in mobile path.

problem is a topic for further research.

Additionally, it would be useful to extend the results of this paper to situations with multiple channels per base station and with more than two bases, and to incorporate the possible non-availability of channels. Work is in progress.

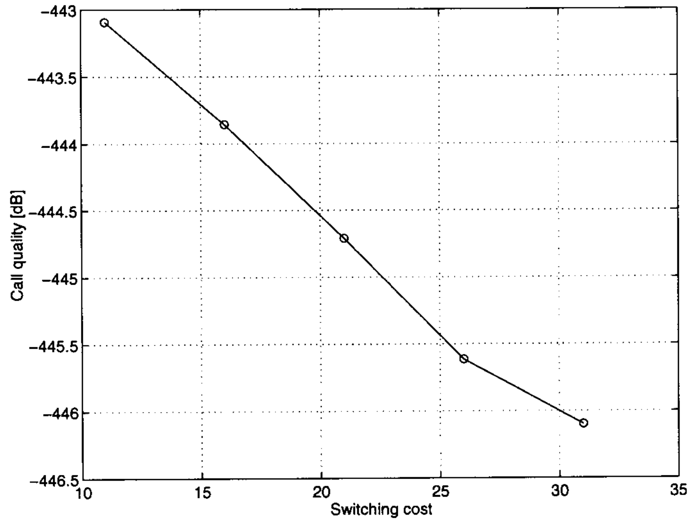

Figure 8: Call quality degrades as the switching cost increases.

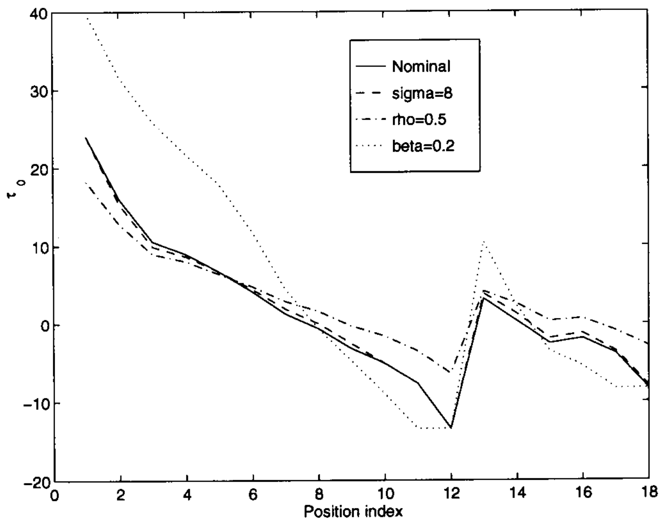

Figure 9: Effect of changing parameters on the optimum threshold.

References

[1] D.P. Bertsekas, *Dynamic Programming: Deterministic and Stochastic Models*, Prentice–Hall, Englewood Cliffs (NJ), 1987.

[2] S.T.S. Chia, "The control of handover initiation in microcells," in *Proceedings of the 41st IEEE Vehicular Conference*, St. Louis (MO), pp. 531–536, May 1991.

[3] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," in *Electron. Lett.*,, vol. 27, no. 23, pp. 2145-2146, Nov. 1991.

[4] W.C.Y. Lee, *Mobile Communications Design Fundamentals*, Second Edition, John Wiley & Sons, New York (NY), 1993.

[5] U. Madhow, M.L. Honig, and K. Steiglitz, "Optimization of wireless resources for personal communications mobility tracking," in *Proceedings of Infocom '94*, Toronto (ON, Canada), pp 577–584 , June 1994.

[6] A. Murase, I.C. Symington, and E. Green, "Handover criterion for macro and microcellular systems," in *Proceedings of the 41st IEEE Vehicular Conference*, St. Louis (MO), pp. 524–530, May 1991.

[7] B. Picinbono, *Random Signals and Systems*, Prentice-Hall, Englewood Cliffs (NJ), 1993.

[8] R. Rezaiifar, *Stochastic Optimization of Handoff in Cellular Networks*, Ph.D. Thesis, Electrical Engineering Department, University of Maryland, College Park (MD), 1995.

[9] S.M. Ross, *Introduction to Stochastic Dynamic Programming*, Academic Press, New York (NY), 1983.

[10] R. Vijayan and J.M. Holtzman, "A model for analyzing handoff algorithms," *IEEE Transactions on Vehicular Technology* **VT–42** (1993), pp. 351–356.

[11] M. Yacoub, *Foundations of Mobile Radio Engineering*, CRC Press, 1993.

[12] P. Whittle, *Optimization Over Time: Dynamic Programming and Stochastic Control, Volume I*, Wiley & Sons, New York (NY), 1982.

[13] P. Whittle, *Optimization Over Time: Dynamic Programming and Stochastic Control, Volume II*, Wiley & Sons, New York (NY), 1982.

A Appendix

A.1 Proof of Proposition 4.1

The setup is that of Sections 4.1 and 4.2. We seek to characterize the conditional distribution of $W_{t+1} = (W^0_{t+1}, W^1_{t+1})$ given (X_t, S_{t+1}). To do so, fix $w_t = (w^0_t, w^1_t)$ in $\mathbb{R}^2$, and s_t, s_{t+1} in E. We readily see from the definitions that

$$\begin{aligned}
&[W_{t+1} \mid W_t = w_t, S_t = s_t, S_{t+1} = s_{t+1}] \\
&=_{st} \left[\begin{matrix} W^0(s_{t+1} - b_0) \\ W^1(s_{t+1} - b_1) \end{matrix} \mid W_t = w_t, S_t = s_t, S_{t+1} = s_{t+1} \right] \\
&=_{st} \left[\begin{matrix} W^0(s_{t+1} - b_0) \\ W^1(s_{t+1} - b_1) \end{matrix} \mid S_t = s_t, S_{t+1} = s_{t+1}, \begin{matrix} W^0(s_t - b_0) = w^0_t \\ W^1(s_t - b_1) = w^1_t \end{matrix} \right] \\
&=_{st} \left[\begin{matrix} W^0(s_{t+1} - b_0) \\ W^1(s_{t+1} - b_1) \end{matrix} \mid \begin{matrix} W^0(s_t - b_0) = w^0_t \\ W^1(s_t - b_1) = w^1_t \end{matrix} \right] \\
&=_{st} \left[\begin{matrix} [W^0(s_{t+1} - b_0) \mid W^0(s_t - b_0) = w^0_t] \\ [W^1(s_{t+1} - b_1) \mid W^1(s_t - b_1) = w^1_t] \end{matrix} \right]
\end{aligned} \tag{A.1}$$

In this last step, the random variables $[W^0(s_{t+1} - b_0) \mid W^0(s_t - b_0) = w^0_t]$ and $[W^1(s_{t+1} - b_1) \mid W^1(s_t - b_1) = w^1_t]$ are taken to be independent by virtue of the independence of the collections of random variables $\{W^i(r),\ r \in \mathbb{R}^2\}$, $i = 0, 1$. This proves Claim **1**.

Because the random variables $\{W^i(r),\ r \in \mathbb{R}^2\}$, $i = 0, 1$, are jointly Gaussian, it is well known [7] that the random variables $[W^i(s_{t+1} - b_i) \mid W^i(s_t - b_i) = w^i_t]$ are also Gaussian with mean c^i_{t+1} and variance C^i_{t+1}, i.e.,

$$[W^i(s_{t+1} - b_i) \mid W^i(s_t - b_i) = w^i_t] \sim \mathcal{N}(c^i_{t+1}, C^i_{t+1}). \tag{A.2}$$

The mean c^i_{t+1} is the conditional expectation of $W^i(s_{t+1} - b_i)$ given $W^i(s_t - b_i) = w^i_t$, and is given by

$$\begin{aligned}
c^i_{t+1} &= \mathbf{E}[W^i(s_{t+1} - b_i) W^i(s_t - b_i)]\, \mathbf{E}[|W^i(s_t - b_i)|^2]^{-1} w^i_t \\
&= \exp\left(-\beta^{-1} \|s_t - s_{t+1}\|\right) w^i_t
\end{aligned} \tag{A.3}$$

Finally, the variance C^i_{t+1} is also the (unconditional) variance

$$C^i_{t+1} = \mathbf{E}[|W^i(s_{t+1} - b_i) - \exp\left(-\beta^{-1} \|s_t - s_{t+1}\|\right) \cdot W^i(s_t - b_i)|^2] \tag{A.4}$$

and Claim **2** is readily established using (A.4) and the enforced independence. ■

A.2 Proof of Proposition 6.1

Fix an element φ in $\mathcal{F}$. For every (x,i) in $E \times \mathbb{R}^2 \times \{0,1\}$ and every s' in E, we have

$$
\begin{aligned}
&\int_{\mathbf{R}^2} |\varphi(s',p',i)| dG(p' \mid x,s') && \text{(A.5)}\\
&= \|\varphi\| \int_{\mathbf{R}^2} (1+K|p'^1 - p'^0|) dG(p' \mid x,s') && \text{(A.6)}\\
&= \|\varphi\| \int_{\mathbf{R}} (1+K|z'|) dF(z' \mid s, p^1-p^0, s') && \text{(A.7)}\\
&\leq \|\varphi\| (1+K(A|p^1-p^0|+B)) && \text{(A.8)}
\end{aligned}
$$

Combining (6.8) and (A.8), we find

$$\|\tilde{T}\varphi\| \leq \sup_{z\in\mathbf{R}} \left\{ \frac{1+K(A|z|+B)}{1+K|z|} \right\} \|\varphi\| \tag{A.9}$$

with the supremum being achieved at $z=0$, and Claim **1**, including (6.10), follows.

Claim **2** is an immediate consequence of Claim **1** once we note that the cost–per–stage c given by (5.3) is indeed an element of $\mathcal{F}$.

To establish Claim **3**, we fix $u = 0,1$ and (x,i) in $E \times \mathbb{R}^2 \times \{0,1\}$, and then consider two elements φ and φ' in $\mathcal{F}$. Obviously,

$$
\begin{aligned}
&|(T_u\varphi)(x,i) - (T_u\varphi')(x,i)|\\
&= (1-\rho)|(\tilde{T}\varphi)(x,i) - (\tilde{T}\varphi')(x,i)|\\
&\leq (1-\rho)\sum_{s'\in E} Q(s;s') \int_{\mathbf{R}^2} |\varphi(s',p',u) - \varphi'(s',p',u)| dG(p' \mid x,s'). && \text{(A.10)}
\end{aligned}
$$

But for each s' in E, we have

$$
\begin{aligned}
&\int_{\mathbf{R}^2} |\varphi(s',p',u) - \varphi'(s',p',u)| dG(p' \mid x,s')\\
&\leq \|\varphi-\varphi'\| \int_{\mathbf{R}^2} (1+K|p'^1 - p'^0|) dG(p' \mid x,s')\\
&= \|\varphi-\varphi'\| \int_{\mathbf{R}} (1+K|z'|) dF(z' \mid s, p^1-p^0, s')\\
&\leq \|\varphi-\varphi'\| (1+K(A|p^1-p^0|+B)). && \text{(A.11)}
\end{aligned}
$$

Combining (A.10) and (A.11) yields

$$|(T_u\varphi)(x,i) - (T_u\varphi')(x,i)| = (1-\rho)\|\varphi-\varphi'\|(1+K(A|p^1-p^0|+B)) \tag{A.12}$$

so that (6.11) holds with constant L given by

$$L \equiv \sup_{z\in\mathbf{R}} \left\{ (1-\rho)\frac{1+K(A|z|+B)}{1+K|z|} \right\}. \tag{A.13}$$

This supremum is achieved at $z=0$ by virtue of the fact that $A \leq 1$, and is therefore given by the expression (6.12) for L. The choice of K implies $0 < L < 1$. ∎

A.3 Proof of Lemma 7.3

We begin with a technical result on non–increasing functions:

Lemma A.1 *Let $h : \mathbb{R} \to \mathbb{R}$ be a continuous function which is non–increasing. For each $C > 0$, the function $H_C : \mathbb{R} \to \mathbb{R}$ defined by*

$$H_C(z) \equiv \min\{C, -z + h(z)\} - \min\{z, C + h(z)\}, \quad z \in \mathbb{R} \tag{A.14}$$

is also continuous and non–increasing.

Proof: The continuity of H_C is obvious. To show that H_C is also non–increasing, we note by direct inspection that

$$\begin{aligned} H_C(z) &= -z + \min\{C + z, h(z)\} - C - \min\{z - C, h(z)\} \\ &= \begin{cases} -z + C & \text{if } z + C \le h(z) \\ -2z + h(z) & \text{if } z - C < h(z) < z + C \\ -z - C & \text{if } h(z) \le z - C. \end{cases} \end{aligned}$$

Next we define the sets A^+ and A^- as:

$$A^+ \equiv \{z \in \mathbb{R} : z + C \le h(z)\}$$

and

$$A^- \equiv \{z \in \mathbb{R} : h(z) \le z - C\}.$$

Using the continuity and non–increasingness of h, we see that these two sets are non–empty, closed and connected subsets which are disjoint. Hence we necessarily have $A^+ = (-\infty, a^+]$ and $A^- = [a^-, \infty)$ with $a^+ < a^-$ (because $C > 0$). To avoid ambiguities, we take $a^+ \equiv \sup A^+$ and $a^- \equiv \inf A^-$. It is then also plain that the set $\{z \in \mathbb{R} : z - C < h(z) < z + C\}$ coincides with the interval (a^+, a^-).

To conclude, we note that H_C is obviously non–increasing on each of the intervals A^+, (a^+, a^-) and A^-, thus on the entire real line by continuity. ■

Take φ in $\mathcal{C} \cap \mathcal{D}$. We first prove that $\widetilde{T}\varphi$ belongs to $\mathcal{C} \cap \mathcal{D}$. By Lemma 7.2 we already know that $\widetilde{T}\varphi$ is an element of $\mathcal{C}$. Next, for each s in E, by (7.3) we have

$$(\Delta\widetilde{T}\varphi)(s, z) \equiv \sum_{s' \in E} Q(s; s') \int_{\mathbb{R}^2} \Delta\varphi(s', z') dF(z' \mid s, z, s') \quad z \in \mathbb{R}. \tag{A.15}$$

Hence, the mapping $z \to \Delta\widetilde{T}\varphi(s, z)$ is seen to be non–increasing once we note from (5.7)-(5.10) that $F(\cdot \mid s, z, s')$ is the distribution of the random variable $\alpha z + \beta + \sqrt{\Theta}U$ with α, β, Θ depending only on s and s', with $\alpha > 0$ and $U \sim \mathcal{N}(0, 1)$.

Next we prove that $T\varphi$ belongs to $\mathcal{C} \cap \mathcal{D}$, For every (s, z, i) in $E \times \mathbb{R}^2 \times \{0, 1\}$, we have

$$\begin{aligned} &T\varphi(s, z, i) \\ &= \min_{u=0,1} \left\{c(z, i, u) + (1 - \rho)(\widetilde{T}\varphi(s, z, u)\right\} \\ &= \min\left\{C + (1 - \rho)(\widetilde{T}\varphi)(s, z, i \oplus 1), (-1)^i z + (1 - \rho)(\widetilde{T}\varphi)(s, z, i)\right\}. \end{aligned}$$

Hence, specializing for $i = 0, 1$, we get

$$\begin{aligned} & T\varphi(s,z,1) \\ &= \min\left\{C + (1-\rho)(\widetilde{T}\varphi)(s,z,0), -z + (1-\rho)(\widetilde{T}\varphi)(s,z,1)\right\} \\ &= (1-\rho)(\widetilde{T}\varphi)(s,z,0) + \min\left\{C, -z + (1-\rho)(\Delta\widetilde{T}\varphi)(s,z)\right\} \end{aligned} \tag{A.16}$$

and

$$\begin{aligned} & T\varphi(s,z,0) \\ &= \min\left\{C + (1-\rho)(\widetilde{T}\varphi)(s,z,1), z + (1-\rho)(\widetilde{T}\varphi)(s,z,0)\right\} \\ &= (1-\rho)(\widetilde{T}\varphi)(s,z,0) + \min\left\{C + (1-\rho)(\Delta\widetilde{T}\varphi)(s,z), z\right\}. \end{aligned} \tag{A.17}$$

Subtracting (A.17) from (A.16) gives,

$$\begin{aligned} (\Delta T\varphi)(s,z) &= \min\left\{C, -z + (1-\rho)(\Delta\widetilde{T}\varphi)(s,z)\right\} - \\ & \min\left\{C + (1-\rho)(\Delta\widetilde{T}\varphi)(s,z), z\right\}. \end{aligned} \tag{A.18}$$

By the first part of the proof, the mapping $z \to (\Delta\widetilde{T}\varphi)(s,z)$ is continuous and non-increasing, and so is $z \to (\Delta T\varphi)(s,z)$ by a straightforward application of Lemma A.1. ■