Integrated Voice/Data Call Admission Control for Wireless DS-CDMA Systems

Sumeetpal Singh, Vikram Krishnamurthy, *Senior Member, IEEE*, and H. Vincent Poor, *Fellow, IEEE*

***Abstract*—This paper addresses the call admission control problem for multiservice wireless code division multiple access (CDMA) cellular systems when the physical layer channel and receiver structure at the base station are taken into account. The call admission problem is formulated as a semi-Markov decision process with constraints on the blocking probabilities and signal-to-interference ratio (SIR). By using recent results in large random matrices, the SIR constraints incorporate linear multiuser detectors and fading channels. We show that the optimal call admission policy can be computed via a linear programming-based algorithm.**

***Index Terms*—Call admission control, CDMA cellular system, fading, semi-Markov decision process.**

I. INTRODUCTION

CODE-DIVISION multiple-access (CDMA) implemented with direct-sequence (DS) spread-spectrum signaling is among the most promising multiplexing technologies for cellular telecommunications services such as personal communications, mobile telephony, and indoor wireless networks. The advantages of DS-CDMA include superior operation in multipath environments, flexibility in the allocation of channels, the ability to operate asynchronously, increased capacity in bursty or fading networks, and the ability to share bandwidth with narrowband communication systems without undue degradation of either system's performance [10].

This paper studies the problem of call admission control (CAC) for an integrated voice/data wireless DS-CDMA system. In the integrated voice/data scenario, users transmit at different bit rates and have different quality-of-service (QoS) requirements, which is usually characterized in terms of a signal-to-interference (SIR) constraint or a bit error probability (BEP) constraint. A multicode (MC) CDMA system is proposed in [4] for integrating users of varying transmission rates. In MC-CDMA, a signature sequence is used to transmit information at a *basic* bit rate. Users that require higher transmission rates use multiple signature sequences in parallel.

When a user desires access to the network for communicating voice or data, the base station (BS) decides whether to admit or block (reject) the new user; this is known as CAC. Assuming that users have stringent SIR requirements, the new user will be blocked if the BS is unable to simultaneously accommodate the SIR requirements of all presently active users and that of the new user.[1] Consider a network in which users are partitioned into K service classes. In addition to satisfying the SIR requirement of all active users, the network also seeks to guarantee that all users in a particular service class will experience a blocking probability of no larger than a particular maximum value. In order to satisfy the SIR and worst-case blocking probability requirements, it will be necessary to admit or block new users as a function of the current profile of active users in the BS. Such an admission mechanism is called a CAC *policy*, and the problem of constructing optimal CAC policies subject to SIR and blocking probability constraints are addressed in this paper.

Existing Approaches: Most existing papers in the literature may be broadly characterized as follows.

i) *Complete sharing:* A policy that accepts a new user if and only if the BS has sufficient capacity to simultaneously accommodate the SIR requirement of all presently active users and the new user is known as *complete sharing* (CS) policy. (The term CS applies when a new user is always offered access to the network provided that there is sufficient bandwidth at the time of request [17].) The CS policy, which is the simplest policy, is proposed and investigated in [6], [11], and [13]. The approaches of [6], [11], and [13] differ in how the BS evaluates the *residual capacity*, which is defined as the additional number of users that can be admitted such that SIR constraints of all active users are satisfied. (Obviously, the residual capacity is a function of the current profile of active users in the BS.) In the case of [6], the residual capacity is defined in terms of the so-called *noise rise condition*; see [6, eq. (5)]. Because the admission and rejection of a new user is based on the SIR constraints, the approaches in [6], [11], and [13] are effectively integrating the network level call admission problem with the performance of the physical layer. However, the shortfalls of [6], [11], and [13] are a) the evaluation of the residual capacity is heuristic as there is no mention of the receiver structure that will be used to demodulate the users received signal at the BS; it is well known that the achieved SIR is maximized by the LMMSE receiver [20] (among the class of linear re-

[1]The term active user here is used for users that have been admitted by the BS.

Manuscript received August 20, 2001; revised February 22, 2002. This work was supported in part by the Centre of Expertise in Networked Decision Systems (CENDS), by an ARC large grant and the ARC special research Centre for Ultra Broadband Information Networks (CUBIN), by the National Science Foundation under Grant CCR-99, and by the New Jersey Center for Wireless Telecommunications. The associate editor coordinating the review of this paper and approving it for publication was Prof. Vahid Tarokh.

S. Singh and V. Krishnamurthy are with the Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, Australia (e-mail: ssss@ee.mu.oz.au; vikram@ee.mu.oz.au).

H. V. Poor is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: poor@ee.princeton.edu).

Publisher Item Identifier S 1053-587X(02)04385-4.

ceivers), and therefore, the residual capacity depends on the receiver structure as well; b) the admission policy used is the simple CS policy, and they do not consider the problem of optimizing the admission policy to minimize the probability of rejecting a new user.

ii) *SMDP approach:* In [5], [17], and [21], the theory of semi-markov decision processes (SMDPs) is used to construct optimal CAC policies. In [17], a circuit switched network is considered, whereas[5] and [21] consider a cellular network. Not only are policies constructed via the SMDP approach are optimal, the SMDP approach can cope with blocking probability constraints. The shortcomings of [5] and [21] are that the admission control problem is treated purely as a network layer problem by ignoring the physical layer aspects such as the CDMA modulation scheme, the channel fading characteristics, receiver structures, SIR, etc.

iii) *Threshold policy:* Several papers have studied the problem of access control for an integrated voice/data CDMA system; see [12], [18], and references within. Access control differs from CAC in the following regard: In access control, the BS admits more users than it has the capacity to accommodate simultaneously. The access control strategy will then regulate the use of the channel by each active user (decide on the bit periods that an active user can transmit) to satisfy a) maximum end-to-end delay for real time voice users and b) SIR constraints for both voice and data users. The CAC policy used in conjunction with access control in [12] and [18] is a simple *threshold policy*, i.e., a new class k user is accepted if the number of active class k users is less than a threshold T_k. Thresholds are then optimized empirically to achieve desired maximum blocking probabilities for the various classes. Although simple to implement, it is well known that threshold policies cannot satisfy blocking probability constraints in general [16]. Additionally, thresholds must be optimized empirically and can perform poorly in practice; see [16] for examples. In any case, if the optimal CAC policy turns out to be a threshold policy, the SMDP approach (being the most general) will identify it.

Contributions: The contributions of this paper are as follows.

- Unlike [5], [6], [11], [13], [17], and [21], we study the interplay between the physical layer interference suppression algorithms (linear multiuser detectors for the CDMA system), the QoS constraints for the various user classes, and the network layer throughput (blocking probability). Specifically, we present a linear programming (LP)-based algorithm for computing the CAC policy that minimizes the blocking probability (or maximizes the network throughput) of a specific service class subject to satisfying constraints on i) the minimum SIR (or maximum BEP) for all active users and ii) the maximum blocking probability for all the other classes of users. In the SIR constraints, we account for LMMSE receivers at the BS and a fading channel using results on the spectrum of large random matrices [8]. By ensuring all active users have a minimum SIR level, not only are we ensuring a maximum value for the BEP, but we also indirectly guarantee a maximum outage probability.[2] As in [5], [17], and [21], the LP-based algorithm for constructing the optimal CAC policy is derived using the the theory of SMDPs.
- In this paper, we address the problem of congestion by generalizing the CAC problem to account for the scenario in which a new data user can be rejected, admitted as an active user (i.e., allocated a signature sequence and allowed to commence transmission immediately), or queued in a finite buffer at the BS. In the work of all previous authors mentioned, when a new user arrives and there is insufficient capacity to support it (i.e., the network is congested), the new user is rejected outright. The option of queuing prevents the new user from being lost to the system, which would be the case if the user was rejected. The queued user is then admitted at a later time when there is sufficient capacity. In numerical examples, we demonstrate significant reductions in the blocking probability of data users when queuing is employed with moderate buffer sizes. Queuing of data users is typical in multiservice wireless networks for data users engaged in nonreal-time services such as e-mail, file transfer, store and forward facsimile, etc.

Finally, we mention that [5] constructs an optimal CAC policy that minimizes the probability of dropping handoff calls. We remark that the formulation of the CAC problem in Section III and the LP algorithm in Section IV-B is general enough to straightforwardly account for the handoff call blocking probability as a performance criterion; see Section V for details.

Limitations: The CAC policy constructed in this paper is optimal when data and voice users arrive according to Poisson processes and have exponential holding (service) times. Under these assumptions, the CAC problem is a SMDP, and one can use a LP algorithm to construct the optimal CAC policy. If we drop the Poisson arrival and exponential holding time assumptions, then the CAC problem is a *generalized semi-Markov process* (GSMP). The optimization of a GSMP is considerably more difficult and is not solved by a LP algorithm, as in the SMDP case. In the GSMP case, our CAC policy can be viewed as suboptimal solution to an otherwise difficult problem.

Notation: For a complex-valued matrix or vector a, a^T, a^*, and a^H, denote the transpose, conjugate, and Hermitian transpose, respectively. I_n denotes the identity matrix of order $n \times n$. $\mathbb{R}_+$ and $\mathbb{Z}_+$ denote the set of non-negative reals and integers, respectively. I is the indicator function, i.e., for a nonempty set A, $I_A(x) = 1$ if $x \in A$, $I_A(x) = 0$ if $x \notin A$. $|A|$ denotes the cardinality of the set A. The function $\delta : \mathbb{R}_+ \rightarrow \{0, 1\}$ is defined by $\delta(x) = 1$ if $x > 0$, $\delta(0) = 0$. For two random variables X and Y, $X \sim Y$ implies that X and Y have the same probability distribution. $\mathbf{P}$ denotes probability, and $\mathbf{E}\{\}$ is the expectation operator. With probability 1 is abbreviated as w.p. 1 .

[2]An outage occurs when the instantaneous SIR falls below the threshold for a prolonged period of time and the outage probability is defined as the percentage of time the instantaneous SIR lies below the threshold [6].

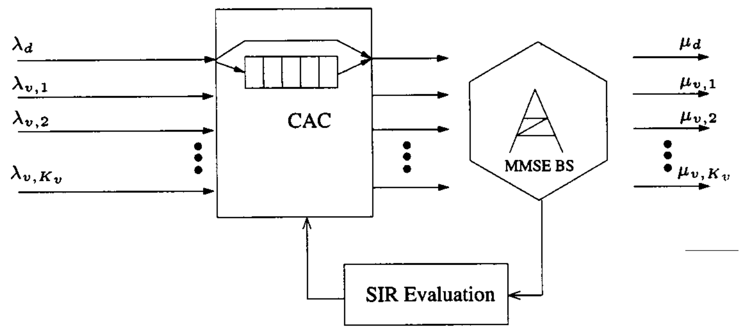


Fig. 1. Call admission for a BS; nonreal-time data users maybe buffered for delayed admission; SIR evaluation is fed back to CAC policy.

II. CAC FOR AN INTEGRATED VOICE/DATA DS-CDMA CELLULAR SYSTEM

In this section, we formulate and solve the CAC problem for the *uplink* (mobile station to BS) of a synchronous DS-CDMA cellular system that integrates both voice and data service, as illustrated in Fig. 1. The signal-model, the receiver structure at the BS, and the corresponding SIR expression are detailed. In Section III, the CAC problem is formulated and solved as a SMDP.

The motivation of the model in Fig. 1 stems from the aim of studying the interplay between the physical layer interference suppression algorithms (linear multiuser detectors), the QoS constraints for the various user classes, and the network layer throughput (blocking probability). The ingredients of the model schematically illustrated in Fig. 1 are as follows.

i) Multiservice user requests to access the wireless DS-CDMA voice/data network.
ii) An admission controller then decides whether to admit the user. The admission controller seeks to optimize the throughput (i.e., minimize the blocking probability) of a specific class of users subject to two types of constraints: a) network level constraint on the throughput of other user classes and b) physical level constraint on the SIR (or BEP) of the various user classes. Additionally, to cope with congestion, there is an option of buffering nonreal-time data users for delayed admission.
iii) If admitted, the user transmits over a fading channel. A LMMSE multiuser detector demodulates each user. The interference suppression capability of the LMMSE multiuser detector is measured by the SIR, which is a surrogate for the BEP.
iv) The evaluated SIR is passed back to the admission controller, which will determine whether or not to admit new users. In this way, the performance of the physical layer interference suppression algorithms (LMMSE detectors) affects the admission of new users.

Details of items i)-iv) follow. Item ii), i.e., Admission Control, is dealt with entirely in Section III .

1) Multiservice User Request: In an integrated voice and data scenario, some of the active users may be engaged in a voice call, and the remaining users will be transmitting data to the BS, i.e., user k is either transmitting voice or data. At connection request, we quantize the transmit power, channel gain, and variance of voice users into a value from the finite set

$$\{(P_{v,i}, \bar{h}_{v,i}, \xi_{v,i}^2), \quad i = 1, \ldots, K_v\}. \tag{1}$$

If the parameters of a voice user are quantized to $(P_{v,i}, \bar{h}_{v,i}, \xi_{v,i}^2)$, the voice user is said to belong to the ith class. In order to base CAC on the performance of the physical layer (e.g., SIR), it is essential that the admission control policy has an indication of transmit power, channel gain, and variance of the user requesting connection; see item iv) in the following for more details. (In [8, Sec.4], the problem of estimating the mean and the variance channel fading process of each user is considered using data that is available during the training period or during the decision-directed mode.) The number of quantization levels in (1) will determine the cardinality of the*state space* of the optimal CAC problem. (See Section III.) Fewer levels will result in a smaller state space and, hence, require fewer computations to compute the optimal CAC policy. For simplicity, we assume only one quantization level for data users requesting service, namely

$$\{(P_d, \bar{h}_d, \xi_d^2)\}. \tag{2}$$

The development that follows may be easily generalized to situation where there are multiple quantization levels as in the voice case; see Section V for more details.

In this paper, we design CAC independently of power control and access control; this is a practically feasible and widely used methodology [6]. The resulting CAC is a static optimization problem since it is only concerned with a user's power and channel parameters at admission request time. Due to time variation in the channel characteristics, these parameters will change after admission. It is the role of access control and power control to dynamically ensure that the SIR requirements of users are satisfied in real time, as is also proposed in [6].

As in [5], we assume that data and class i voice calls are generated according to homogeneous Poisson processes with intensities

$$\lambda_d \quad \text{and} \quad \lambda_{v,i} \tag{3}$$

respectively. Furthermore, the duration of a data and a class i voice call $i \in \{1, \ldots, K_v\}$ are exponential random variables with means

$$\frac{1}{\mu_d} \text{ and } \frac{1}{\mu_{v,i}} \tag{4}$$

respectively.

2) Fading DS-CDMA Wireless Channel and LMMSE Detector: The signal model we use is the "standard" model for a synchronous (or asynchronous) DS-CDMA system in a fading environment [8], [14]. Consider the uplink of a synchronous K-user DS-CDMA communication system. Assume that this system transmits binary symbols through an additive white Gaussian noise channel in a single-path (flat) fading environment. The chip-sampled discrete-time model for the received baseband signal (at the BS) during the nth bit interval is

$$r(n) = \sqrt{P_1} b_1(n) h_1(n) s_1 + \sum_{k=2}^{K} \sqrt{P_k} b_k(n) h_k(n) s_k + \sigma w(n). \quad (5)$$

There K active users P_k, $b_k(n)$ and $h_k(n)$ denote the transmit power, the nth transmitted bit, and channel gain for the kth user, respectively; $s_k \in \mathbb{R}^N$ is the signature sequence for the kth user, where N is the processing gain (spreading gain); $w(n)$ is additive noise, and $\sigma > 0$ determines its variance. It is assumed that $\{b_k(n)\}_{n \in \mathbb{Z}_+}$ is a sequence if independent and identically distributed (i.i.d.) equiprobable ± 1 random variables and that $\{h_k(n)\}_{n \in \mathbb{Z}_+}$ is a complex-valued random process satisfying $\mathbf{E}\{h_k(n)\} = \bar{h}_k$ and $\mathbf{E}\{h_k(n) h_k^*(n)\} - \mathbf{E}\{h_k(n)\}(\mathbf{E}\{h_k(n)\})^* = \xi_k^2$. The additive noise $\{w(n)\}_{n \in \mathbb{Z}_+}$ is a circularly symmetric complex white Gaussian random process with $\mathbf{E}\{w(n)\} = [0, \ldots, 0]^T$ and $\mathbf{E}\{w(n) w^*(n)\} = I_N$. It is further assumed that the stochastic processes $\{b_k(n)\}_{n \in \mathbb{Z}_+}$, $\{h_k(n)\}_{n \in \mathbb{Z}_+}$, $k = 1, \ldots K$ and $\{w(n)\}_{n \in \mathbb{Z}_+}$ are mutually independent. As in [8], we assume that the path delays induced by the fading channels are negligible compared with the bit duration NT_c, where T_c is the chip-period, i.e., intersymbol-interference is negligible. Extension to the intersymbol-interference case is possible and is remarked in Section V.

Assume that user 1 is the user of interest and that we wish to "recover" the transmitted bit $b_1(n)$ from $r(n)$ defined in (5). A linear demodulator for user 1 is a vector $c_1 \in \mathbb{C}^N$ (N-dimensional column vector with complex elements) that is used to obtain an estimate of the transmitted bit as follows: $\hat{b}_1(n) = \text{sgn}(Re[c_1^H r_n])$. For user 1, the LMMSE detector [14] chooses the vector $c_1^\star \in \mathbb{C}^N$ that minimizes the mean squared error $\mathbf{E}\{(b_1(n) - c_1^H r(n))(b_1(n) - c_1^H r(n))^*\}$, which is given by

$$c_1^\star = (\mathbf{E}\{r(n) r^H(n)\})^{-1} \mathbf{E}\{b_1(n) r(n)\}. \quad (6)$$

(In this section, when the expectation operator is invoked, we always condition on the information that is assumed available at the LMMSE receiver; we abbreviate $\mathbf{E}\{.|\mathcal{I}\}$ to $\mathbf{E}\{.\}$, where $\mathcal{I}$ represents knowledge of the transmit power, the mean and variance of the channel, and the signature sequence of all multiple-access users.)

Write the parameters of the signal model (5) compactly as follows:

$$\begin{aligned} D_1 =& \text{diag}(P_2 \mathbf{E}\{h_2(n) h_2^*(n)\}, \ldots, P_K \mathbf{E}\{h_K(n) h_K^*(n)\}) \\ D =& [P_1 \mathbf{E}\{h_1(n) h_1^*(n)\}, D_1], \; S_1 = [s_2, \ldots, s_K] \\ S =& [s_1, S_1]. \end{aligned}$$

Using the independence assumptions on the various stochastic processes in the previous paragraph, as well as the first- and second-order statistics given therein, we have

$$c_1^\star = \sqrt{P_1} \bar{h}_1 (SDS^T + \sigma^2 I_N)^{-1} s_1. \quad (7)$$

3) Evaluated SIR: The BEP is the main QoS (performance) measure in wireless networks. In this paper, BEP constraints are accounted for by introducing SIR constraints. It is well known that SIR is a surrogate measure for the BEP; intuitively, this is because the BEP is degraded by the interference introduced by the multiaccess users and the background channel noise. One a more technical note, it has been established that in a "large system" (K and N in (5) are large), the BEP monotonically decreases as the SIR increases [22, Th. 3.3].

Let SIR_1 denote signal-to-interference (SIR) ratio for the estimate $c_1^{\star H} r(n)$, with $c_1^\star$ given in (7). Then, as shown in [8]

$$\text{SIR}_1 = P_1 |\bar{h}_1|^2 s_1^T (S_1 D_1 S_1^T + \sigma^2 I_N + P_1 \xi_1^2 s_1 s_1^T)^{-1} s_1. \quad (8)$$

When a new user requests to be admitted into the network, the BS must ensure that the SIR constraints of all presently active users and that of the new user can be simultaneously satisfied. Note that the SIR (8), amongst other factors, is a function of the signature sequence of the user of interest and that the remaining $K-1$ active users. By recourse to the exact expression for the SIR (8), it is impossible to design an optimal CAC algorithm that is *computationally feasible* and, hence, of practical interest. This is because the definition of the state space of the CAC problem (see Section III) must include *the set of all possible signature sequences*. (This is the set from which users transmitting in the cell are allocated signature sequences from.) As in [8], [9], [20], and [22], we assume that signature sequences of the K users (5) are *randomly and independently* chosen on admission; this implies that the signature sequence for user k can be modeled as $s_k = (1/\sqrt{N})[\nu_{k1}, \ldots, \nu_{kN}]^T$, where ν_{kl}'s are i.i.d. with mean zero and variance 1. [A candidate for ν_{kl} could be the equiprobable ± 1 random variable. The normalization $(1/\sqrt{N})$ ensures $\mathbf{E}\{s_k^T s_k\} = 1$.] Under this assumption, it was shown [8] that the SIR can be closely approximated by an expression that only depends on the transmit powers of all active users as well as the first- and second-order statistics of the channel gain processes

$$\text{SIR}_1 \approx \widehat{SIR}_1 \triangleq \frac{P_1 |\bar{h}_1|^2 \beta}{1 + P_1 \xi_1^2 \beta} \quad (9)$$

where β is the unique fixed point in (0, ∞) that satisfies

$$\beta = \left[\sigma^2 + \frac{1}{N} \sum_{k=2}^{K} I\left(P_k(\xi_k^2 + |\bar{h}_k|^2), \beta \right) \right]^{-1} \quad (10)$$

and

$$I(p, \beta) \triangleq \frac{p}{1 + p\beta}. \quad (11)$$

For the special case when $K = 1$, i.e., only the user 1 is active, SIR_1 is approximated by

$$\widehat{\text{SIR}}_1 \triangleq \frac{\frac{P_1 |\bar{h}_1|^2}{\sigma^2}}{1 + \frac{P_1 \xi_1^2}{\sigma^2}}. \quad (12)$$

For the case when the channel mean is 1 and the variance is 0, i.e., there is no fading, [20] presents numerical examples that verify the accuracy of the above approximations; see also [9] for an analysis of the SIR based on the spectrum of large random matrices. In [8, Th. 1], a convergence result is presented that shows $|\text{SIR}_1 - \widehat{\text{SIR}}_1|$ converges (almost surely) to 0 as N is increased to infinity, whereas the ratio K/N is held fixed at some positive value. This convergence result, loosely speaking, states that the approximation is accurate for "large system," i.e., when K and N are large or for a system with large spreading gain N.

The expression for the SIR in (9) is independent of the signature sequences and depends only on the transmit power, the channel mean, and the channel variance of all active users. This approximate expression will be used to construct the state space of the optimal CAC problem in Section III. If user 1 has a specified SIR lower threshold, say $\underline{\text{SIR}}_1$, then as shown in the following proposition, one need not solve for the fixed point β of (9) to verify $\widehat{\text{SIR}}_1 \geq \underline{\text{SIR}}_1$.

Proposition II.1: Suppose $\underline{\text{SIR}}_1$ is a specific lower threshold for $\widehat{\text{SIR}}_1$ defined in (9). Then, $\widehat{\text{SIR}}_1 \geq \underline{\text{SIR}}_1$ if and only if

$$\left[\sigma^2 + \frac{1}{N}\sum_{k=2}^{K} I\left(P_k(\xi_k^2 + |\bar{h}_k|^2), \gamma\right)\right]^{-1} \geq \gamma \qquad (13)$$

where

$$\gamma = \frac{\underline{\text{SIR}}_1}{P_1|\bar{h}_1|^2 - \underline{\text{SIR}}_1 P_1 \xi_1^2}. \qquad (14)$$

III. Formulation of the CAC Problem as a SMDP

The aim of this section is to formulate the CAC problem for the uplink (see Fig. 1) as an SMDP. The methodology presented here still applies if admission control is to be performed based on both the *uplink and downlink* capacity or the *downlink* capacity alone.

The SMDP formulation below proceeds in three steps.

1) First, a discrete-valued (finite) state space for the profile of active users in the network is specified. The SIR constraints are incorporated by truncating the state space to those points that satisfy the SIR constraints. A SMDP is then defined over this truncated state space. From an implementation point of view, constructing the truncated state space can be difficult as it would involve exhaustive enumeration. We show that the SIR constraints induce a convex SMDP state space and give that a simple procedure exists for constructing it in Section V.

2) The actions and the state dynamics of the SMDP are then defined. The arrival process for each user class is assumed to be a continuous-time homogeneous Poisson process. The duration of data and voice calls are assumed to be exponentially distributed. The class of admissible CAC policies are then defined.

3) The performance criterion for minimizing the blocking probability is specified. We show i) that the blocking probability is captured by an additive cost function (via the PASTA theorem [7]) and ii) that all CAC policies within the admissible are *unichain*. By virtue of i) and ii), the standard approach for solving a SMDP via the linear programming method applies [1, Ch. 5]. In particular, one may write a LP whose solution is the optimal CAC policy. Blocking probability constraints for the various user classes are accommodated by merely adding additional linear constraints to this LP.

An SMDP is characterized by the following ingredients.

1) *State Space:* Let X denote the *state space* and $x(t) \in X \subset \mathbb{Z}_+^{2+K_v}$ the *state* of the BS at time t, where $t \in \mathbb{R}_+$. The state vector is given by

$$x(t) = [x_d(t), x_b(t), x_{v,1}(t), x_{v,2}(t), \ldots, x_{v,K_v}(t)]^T \qquad (15)$$

where $x_d(t)$ denotes the number of active data users, $x_b(t)$ denotes the number of data users in the buffer, and $x_{v,i}$, $i = 1, \ldots, K_v$ denotes the number of active class i voice users, where K_v is the total number of voice classes. X is a finite set [see (17)]. We assume the buffer for data users is a finite buffer of length B_d. The vector

$$\beta = [\beta_d, \beta_{v,1}, \beta_{v,2}, \ldots, \beta_{v,K_v}]^T \qquad (16)$$

characterizes the minimum SIR that all active data and voice calls of each class type must satisfy. For the current state $x(t) = x$, let $\Psi_d(x)$ denote the SIR value for all active data users, and let $\Psi_{v,i}(x)$ denotesthe SIR value for all active class i voice users. By convention, set $\Psi_d(x) = \infty$ ($\Psi_{v,i}(x) = \infty$) if $x = [x_d, x_b, x_{v,1}, x_{v,2}, \ldots, x_{v,K_v}]^T$ is such that $x_d = 0$ ($x_{v,i} = 0$). (These *SIR functions* are defined in Section IV-A.) The state space X is defined as

$$\begin{aligned} X = \{x = &[x_d, x_b, x_{v,1}, \ldots, x_{v,K_v}]^T \in \mathbb{Z}_+{}^{2+K_v}: \\ &x_b \leq B_d,\ \Psi_d(x) \geq \beta_d, \\ \Psi_{v,i}(x) &\geq \beta_{v,i}, i = 1, \ldots, K_v\}. \end{aligned} \qquad (17)$$

X comprises of the data buffer states together with all combinations of data and voice users that satisfy the SIR thresholds (16). For the SIR functions defined in the following, X is a finite set. Additionally, since the arrival and departure of calls are random, and $\{x(t)\}_{t\in\mathbb{R}_+}$ is a finite-state stochastic process.

2) *Decision Epochs, Actions, and State Dynamics:* When an arriving data or voice user desires to be admitted into the system, the BS will make a decision as to whether or not to grant admission. Thus, a "natural" definition for the decision epochs are the *arrival instances* of the data and voice users, as is done in [15, Ch. 11]. However, as in [17], we define the decision epochs as the instances when the stochastic process $\{x(t)\}_{t\in\mathbb{R}_+}$ changes state, i.e., arrivals and departure are taken into account. Formally, let

$$t_k \text{ denote the } k\text{-th transition time of } \{x(t)\}_{t\in\mathbb{R}_+}. \qquad (18)$$

By convention, set $t_0 = 0$. The decision epochs are taken to be the instances t_k, $k = 0, 1, 2, \ldots$ At each decision epoch, an admit or block decision is made for each possible type of arrival that may occur in the time interval $(t_k, t_{k+1}]$. These decisions are collectively referred to as an *action*. The set of all possible actions (action space) A is defined as

$$A = \left\{a = (a_d, a_b, a_{v,1}, a_{v,2}, \ldots, a_{v,K_v}) : a \in \{0,1\}^{2+K_v}\right\} \qquad (19)$$

where the actions $a_d, a_b, a_{v,1}, \ldots, a_{v,K_v}$ are defined in te following. Assuming $x(t_k) = x$, the action at decision epoch t_k, which is denoted by $a(t_k)$, must be selected from the state-dependent subset of A, i.e.,

$$a(t_k) \in A_x \subseteq A \text{ if } x(t_k) = x. \tag{20}$$

Action $a(t_k) = (a_d, a_b, a_{v,1}, \ldots, a_{v,K_v}) \in A$ is interpreted as follows.

- If $a_d = 1$, a data user that arrives in the interval $(t_k, t_{k+1}]$ is admitted as an active user. Otherwise, it is placed in the buffer, provided the buffer is not full. If the buffer is full, the user is blocked.
- If $a_b = 0$, no data users that are queued in the buffer are made active in the interval $(t_k, t_{k+1}]$. If $a_b = 1$ and the buffer is not empty, admit the data user at the head of the buffer as an active user if $t_k + \epsilon \in (t_k, t_{k+1}]$, where $\epsilon \geq 0$ is a randomly generated delay. ϵ is called the *buffered data admit latency* and is intentionally introduced. Note that ϵ need not be random and could be a (arbitrarily small but positive) deterministic delay. See Remark III.1.
- If $a_{v,i} = 1$, admit a class i voice user that arrives in the interval $(t_k, t_{k+1}]$. Otherwise, the user is blocked.

Note that an arriving data or voice user that is blocked does not cause a state transition in the process $\{x(t)\}_{t\in\mathbb{R}_+}$(15); a state transition occurs when an arriving data or voice user is admitted, a data user in the buffer is made active, or an existing active user departs the system.

For a given state $x \in X$, define the admissible action space $A_x \subseteq A$ as follows:

$$\begin{aligned} A_x = \{a \in A :\ & a_d = 0 \text{ if } x_b > 0 \text{ or } x + e_1 \notin X \\ & a_b = 0 \text{ if } x_b = 0 \text{ or } x + e_1 - e_2 \notin X \\ & a_{v,i} = 0 \text{ if } x + e_{i+2} \notin X, \text{ for } i = 1, \ldots, K_v\} \end{aligned} \tag{21}$$

where $e_i \in \mathbb{R}^{2+K_v}$ denotes the vector of all zeros except for the ith component, which is 1. Essentially, A_x is composed of all those actions in A that do not result in a transition to a state $y \notin X$(17). Additionally, the restriction $a_d = 0$ if $x_b > 0$ ensures data users are made active on a first-come-first-serve basis. (This restriction is not necessary but is imposed for "fairness.") Although not explicitly stated in (21), action $(0, 0, \ldots, 0)$ is excluded while in state $[0, B_d, 0, \ldots, 0]^T$. Clearly, such an action is undesirable in the sense that new users are never admitted into the system and, therefore, may be excluded.

The state dynamics of an SMDP [1], [15] are completely specified by stating the transition distributions $Q_{xy}(\tau, a) \triangleq \mathbf{P}(t_{k+1} - t_k \leq \tau, x(t_{k+1}) = y | x(t_k) = x, a(t_k) = a)$ for all $x, y \in X$, $a \in A_x$, and $\tau \in \mathbb{R}_+$; t_k are defined in (18). However, from the point of view of the solution methodology, it is sufficient to deal with the transition probabilities of the so-called *embedded chain* $p_{xy}(a)$ and the expected sojourn time (holding time) $\tau_x(a)$ for each state action pair instead [1, Ch. 5]:

$$\begin{aligned} p_{xy}(a) &\triangleq \mathbf{P}(x(t_{k+1}) \\ &= y | x(t_k) = x, a(t_k) = a) \end{aligned} \tag{22}$$

$$\begin{aligned} \tau_x(a) &\triangleq \mathbf{E}\{t_{k+1} - t_k | x(t_k) \\ &= x, a(t_k) = a\}. \end{aligned} \tag{23}$$

Let the buffered data admit latency be generated according to the exponential distribution, i.e.,

$$\epsilon \sim \exp\left(\frac{1}{\lambda_b}\right) \tag{24}$$

where $\lambda_b > 0$. (See Remark III.1.) Then, for each $x, y \in X$ (15) and $a \in Ax$ (21), the quantities in (22) and (23) can be expressed as

$$\begin{aligned} \tau_x(a) = \Big[& \lambda_d a_d + \lambda_d (1 - a_d) \delta(B_d - x_b) + \delta(x_b) a_b \lambda_b \\ & + x_d \mu_d + \sum_{i=1}^{K_v} (a_{v,i} \lambda_{v,i} + x_{v,i} \mu_{v,i}) \Big]^{-1} \end{aligned} \tag{25}$$

$$p_{xy}(a) = \begin{cases} \lambda_d a_d \tau_x(a), & \text{if } y = x + e_1 \\ \mu_d x_d \tau_x(a), & \text{if } y = x - e_1 \\ \lambda_d (1 - a_d) \delta(B_d - x_b) \tau_x(a), & \text{if } y = x + e_2 \\ \lambda_{v,i} a_{v,i} \tau_x(a), & \text{if } y = x + e_i \\ & i = 3, \ldots, 2 + K_v \\ \mu_{v,i} x_{v,i} \tau_x(a), & \text{if } y = x - e_i \\ & i = 3, \ldots, 2 + K_v \\ \delta(x_b) a_b \lambda_b \tau_x(a), & \text{if } y = x + e_1 - e_2 \\ 0, & \text{otherwise.} \end{cases} \tag{26}$$

Note that $y = x + e_i$ (respectively, $y = x - e_i$) corresponds to an arrival of a new (respectively, departure of an existing) user whose type is made specific from the value of i. $y = x + e_1 - e_2$ is the successor state that results when a data user from the buffer is made active. The expressions in (25) and (26) can be explained as follows: If $x(t_k) = x$ and $a(t_k) = a$, then the new state $x(t_{k+1})$ and holding time in the current state $t_{k+1} - t_k$ is determined by a superposition of mutually independent Poisson processes. The resulting process consists of one departure process with rate $\mu_{v,i}$ for each active class i user, one departure process with rate μ_d for each active data user, an arrival process with rate $\lambda_{v,i}$ if action a admits a class i voice user (i.e., a is such that $a_{v,i} = 1$), an arrival process with rate λ_d if $a_d = 1$ or the buffer is not full and finally, and a Poisson process with rate λ_b if action a admits a data user from the buffer as an active user. In addition, note that if a stochastic process $\{Z(t)\}_{t\in\mathbb{R}_+}$ is a superposition of n independent Poisson processes with rates λ_i, $i = 1, \ldots, n$, then i) the expected interarrival time of the process $\{Z(t)\}$ is $1/(\lambda_1 + \lambda_2 + \cdots + \lambda_n)$, and ii) the probability that the kth event of $\{Z(t)\}$ is generated by the jth Poisson process is $\lambda_j / (\sum_i \lambda_i)$ [3, Sec.6.7.3]. For a more rigorous derivation of (25) and (26), one can write the CAC problem as a *stochastic timed automata with a Poisson clock structure* and use the result in [3, eq. (6.50)] for (25) and the result in [3, eq. (6.68)] for (26).

3) *Policy:* Let $\mathcal{U}$ denote the class of *admissible* CAC policies, which is defined as follows:

$$\mathcal{U} = \{\tilde{u} : X \to A \mid \tilde{u}(x) \in A_x, \ \forall x \in X\}. \tag{27}$$

Note that the definition in (27) includes threshold policies as a special case [16]. Given any $\tilde{u} \in \mathcal{U}$, call admission is performed as follows: For the interval $(t_k, t_{k+1}]$, the action (20) chosen is $a(t_k) = \tilde{u}(x(t_k))$.

4) *Performance Criterion:* We consider the so-called *average cost* criterion[1, Ch. 5]. Let

$$c : X \times A \to \mathbb{R} \tag{28}$$

be uniformly bounded. The performance criterion considered, which includes data user blocking probability as a special case is given as follows: For any $\tilde{u} \in \mathcal{U}$ and $x_0 \in X$, define

$$J_{\tilde{u}}(x_0) = \lim_{T\to\infty} \frac{1}{T} \mathbf{E}\left\{ \int_0^T c(x(t), a(t))\, dt \right\} \tag{29}$$

where $a(t) = \tilde{u}(x(t_k))$ for $t \in [t_k, t_{k+1})$, $x(0) = x_0$. The aim is to compute the optimal policy $\tilde{u}^* \in \mathcal{U}$ that satisfies

$$J_{\tilde{u}^*}(x_0) = \min_{\tilde{u}\in\mathcal{U}} J_{\tilde{u}}(x_0) \quad \text{for all } x_0 \in X \tag{30}$$

i.e., $\tilde{u}^*$ has the minimum cost for all initial states. The limit in (29) exists for any $x_0 \in X$ and $\tilde{u} \in \mathcal{U}$ and may be verified using the theory of *renewal reward processes* [19, p. 219]. In fact, for any $\tilde{u} \in \mathcal{U}$, the limit in (29) is independent of x_0 (see Proposition III.1.). The independence of the limit, which holds because all the policies in $\mathcal{U}$ are *unichain*, is the basis the LP solution methodology in Section IV-B [1] (see the proof of Proposition III.1. for details). A policy $\tilde{u}$ is said to be unichain if the corresponding *embedded Markov chain* has no two disjoint closed sets of states. The embedded Markov chain corresponding to $\tilde{u}$ is the controlled discrete time Markov chain $\{x(t_k)\}_{k\in\mathbb{Z}_+}$ [t_k defined in (18)] with transition probabilities $p_{xy}(\tilde{u}(x))$(22). A nonempty set $\tilde{X} \subseteq X$ is said to be closed if $p_{xy}(\tilde{u}(x)) = 0$ when $x \in \tilde{X}$ and $y \notin \tilde{X}$; see also [1] and [19].

We now state the form of the function c such that the data-blocking probability corresponding to a CAC policy $\tilde{u}$ is given by $J_{\tilde{u}}(x_0)$ in (29). Consider an arbitrary CAC policy $\tilde{u} \in \mathcal{U}$ and an initial state $x_0 \in X$ for the process $\{x(t)\}_{t\in\mathbb{R}_+}$. Let $\{N_d(t)\}_{t\in\mathbb{R}_+}$ denote the *counting process* that counts the number of data user arrivals in the interval $(0, t]$. Let $\bar{N}_d(t)$ denote the number of data users that are blocked in the interval $(0, t]$.[3] The empirical data-blocking probability corresponding to $\tilde{u}$ and x_0, $B^d_{\tilde{u}}(x_0)$ is "defined" as

$$B^d_{\tilde{u}}(x_0) = \lim_{T\to\infty} \frac{\bar{N}_d(T)}{N_d(T)}. \tag{31}$$

At first sight, $B^d_{\tilde{u}}(x_0)$ does not appear to be of the form in (29), i.e., an expected value of the sample path average of $c(x(t), a(t))$. The aim of the following proposition is to show the following ergodicity type result: minimizing the average cost performance criterion $J_{\tilde{u}}(x_0)$ in (29) when $c(x, a) = (1 - a_d)(1 - I_X(x + e_2))$ is equivalent to minimizing the blocking probability $B^d_{\tilde{u}}(x_0)$. (The Poisson assumption for the arrival process of data users is crucial for this result to be true; see the proof of Proposition III.1. for details.) Given this equivalence, we can then use any of the several available algorithms that solve (30) to construct the CAC policy that yields the minimum data (or class i voice) blocking probability.

[3] Obviously, $\bar{N}_d(t)$ is a function of $\tilde{u}$ and the initial state x_0; the dependency is not made explicit in the notation, however.

Proposition III.1 (Blocking Probability as Average Cost):

i) Data: Define c in (28) to be

$$c(x, a) = (1 - a_d)(1 - I_X(x + e_2)), \quad \forall x \in X, \forall a \in A \tag{32}$$

where a_d is defined in (19), and e_2 is defined in (26). Then, for any $\tilde{u} \in \mathcal{U}$, $B^d_{\tilde{u}}(x_0) = J_{\tilde{u}}(x_0)$ w.p. 1 for all $x_0 \in X$, where $J_{\tilde{u}}(x_0)$ is defined in (29) and $B^d_{\tilde{u}}(x_0)$ is defined in (31). Moreover, for all $\tilde{u} \in \mathcal{U}$, $J_{\tilde{u}}(x_0)$ is independent of x_0.

ii) Voice: Similar to the data case, setting

$$c(x, a) = 1 - a_{v,i} \tag{33}$$

yields the minimum blocking probability for class i voice users.

iii) Multiservice: Choosing

$$c(x, a) = \nu_1(1 - a_d)(1 - I_X(x + e_2)) + \sum_{i=1}^{K_v} \nu_{i+1}(1 - a_{v,i}) \tag{34}$$

for some weights $\nu_i \in \mathbb{R}_+$ yields the policy that minimizes a weighted sum of the data and voice classes blocking probabilities.

(All proofs are in the Appendix.)

Remark III.1 (Buffered Data Admit Latency (24)): To ensure $\tau_x(a) > 0$ for all $x \in X$ and $a \in A_x$, the data user from the buffer has to be admitted after some nonzero delay with positive probability. (The condition $\tau_x(a) > 0$ is needed to be able to apply the LP solution methodology in Section IV [1, Ch. 5].) ϵ could even be selected to be an arbitrary positive deterministic delay. Everything remains the same, except for changes to the particular expressions for $p_{xy}(a)$(26) and $\tau_x(a)$(25).

Remark III.2 [Semi-Markov (SM) Property]: If the data buffer admit latency ϵ is an exponential random variable (24), then the length of time (holding time) between two decision epochs $t_{k+1} - t_k$(18) is an exponential random variable as well with rate depending on the state $x(t_k)$ and action $a(t_k)$; hence, we have the expression in (25). In this case, the CAC problem is a special case of a SMDP, which is a *continuous time Markov decision process* (CMDP). When ϵ is a positive deterministic constant, the length of time between two decision epochs $t_{k+1} - t_k$ is no longer an exponential random variable, and the CAC problem is a SMDP. See [1] and [15] for background theory on SMDPs.

IV. SIR Constraints and the LP Solution Methodology

A. SIR Constraints

The SIR functions Ψ_d and $\Psi_{v,i}$ that were used to define the state space in (17) are taken from the approximate expression for the SIR given in (9). Define the functions $f_d : \mathbb{Z}_+{}^{2+K_v} \times \mathbb{R} \to \mathbb{R}$, $f^i_v : \mathbb{Z}_+{}^{2+K_v} \times \mathbb{R} \to \mathbb{R}$, $i = 1, \ldots, K_v$ as follows:

$$f_d(x, \beta) = \Bigg[\sigma^2 + \frac{x_d - 1}{N} I(P_d \xi_d^2 + P_d|\bar{h}_d|^2, \beta) + \sum_{i=1}^{K_v} \frac{x_{v,i}}{N} I(P_{v,i}\xi_{v,i}^2 + P_{v,i}|\bar{h}_{v,i}|^2, \beta)\Bigg]^{-1} \tag{35}$$

and

$$f_v^i(x,\beta)=\Big[\sigma^2+\frac{x_d}{N}I(P_d\xi_d^2+P_d|\bar{h}_1|^2,\beta)+\frac{x_{v,i}-1}{N}I(P_{v,i}\xi_{v,i}^2+P_{v,i}|\bar{h}_{v,i}|^2,\beta)+\sum_{j=1,j\neq i}^{K_v}\frac{x_{v,j}}{N}I(P_{v,j}\xi_{v,j}^2+P_{v,j}|\bar{h}_{v,j}|^2,\beta)\Big]^{-1} \tag{36}$$

where N is the processing gain, and I is defined in (9). Equations (35) and (36) are derived from (9) by taking into account the structure of the multiaccess interference for the integrated voice and data scenario, i.e., an active data user suffers interference from the remaining x_d-1 active data users and $\sum_{i=1}^{K_v}x_{v,i}$ active voice users. Now, the SIR functions are defined as follows:

$$\Psi_d(x)=\begin{cases}\infty, & \text{if } x_d=0\\ \frac{P_d|\bar{h}_d|^2}{\sigma^2+P_d\xi_d^2}, & \text{if } x=e_1+me_2\\ \frac{P_d|\bar{h}_d|^2\beta}{1+P_d\xi_d^2\beta}, & \beta>0:\beta=f_d(x,\beta) \text{ else}\end{cases} \tag{37}$$

$$\Psi_{v,i}(x)=\begin{cases}\infty, & \text{if } x_{v,i}=0\\ \frac{P_{v,i}|\bar{h}_{v,i}|^2}{\sigma^2+P_{v,i}\xi_{v,i}^2}, & \text{if } x=e_{2+i}+me_2\\ \frac{P_{v,i}|\bar{h}_{v,i}|^2\beta}{1+P_{v,i}\xi_{v,i}^2\beta}, & \beta>0:\beta=f_v^i(x,\beta) \text{ else}\end{cases} \tag{38}$$

where m is any non-negative integer such that $m\leq B_d$. Note that in (37), β is the unique positive fixed point that satisfies $\beta=f_d(x,\beta)$ and similarly for β in (38). In the above definition, $\Psi_d(e_1+me_2)=\Psi_d(e_1)$ since users in the buffer do not degrade the SIR. In constructing the state space X(17), for a given $x\in\mathbb{Z}_+^{2+K_v}$, one needs to verify that $\Psi_d(x)\geq\beta_d$, $\Psi_{v,i}(x)\geq\beta_{v,1}$ and so on. There is no need to compute the fixed points as indicated in (37) and (38). As is shown in Proposition II.1, one merely verifies the inequality in (13).

The following proposition states that the state space X induced by the SIR functions Ψ_d and $\Psi_{v,i}$ is convex.

Proposition IV.1: X in (17) satisfies the following two properties for any $x\in\mathbb{Z}_+^{2+K_v}$.

i) If $x+e_j\in X$ for some j, then $x\in X$;
ii) If $x\notin X$, then $x+e_j\notin X$ for all j

where e_j is defined in (21).

Property i) of Proposition IV.1 implies that if all active users satisfy their SIR requirements, then the departure of any active user results in the remaining users still satisfying their SIR requirements. Property ii) implies that if the current profile of active users does not satisfies the SIR constraints, then the same will be true if a new user is made active. These properties make the task of constructing X in (17) simple; see Section VI for details.

B. Constructing the Optimal CAC Policy $\tilde{u}^*$ - LP Formulation

$\tilde{u}^*$ Without Blocking Probability Constraints: SMDPs are usually analyzed and solved within the framework of discrete-time average cost Markov decision processes; this is achieved by a process called*uniformization*. (See [1] for a detailed discussion.) Essentially, one defines and solves an "equivalent" discrete time-average cost Markov decision process (MDP) in lieu of the original SMDP; the discrete time MDP is equivalent in the sense that the solution to its Bellman equation coincides with that of the Bellman equation for the original SMDP. Thus, algorithms for solving discrete time MDPs, such as value iteration, policy iteration, and linear programming, are applicable to SMDPs as well. We will use the LP approach in the following.[4]

The optimal CAC policy $\tilde{u}^*$(30) is obtained by solving the following LP.

$$\begin{aligned}&\min_{z_{xa}\geq 0,\ x\in X, a\in A_x}\sum_{x\in X}\sum_{a\in A_x}c(x,a)\tau_x(a)z_{xa}\\ &\text{subject to}\sum_{a\in A_y}z_{ya}-\sum_{x\in X}\sum_{a\in A_x}p_{xy}(a)z_{xa}=0\qquad y\in X\\ &\sum_{x\in X}\sum_{a\in A_x}\tau_x(a)z_{xa}=1.\end{aligned} \tag{39}$$

The decision variables are z_{xa}, $x\in X$, and $a\in A_x$. Let z_{xa}^* denote the optimal solution to the above LP. The optimal policy $\tilde{u}^*$ is then constructed as follows [19, p. 224]: For each $x\in X$, $\tilde{u}^*(x)\triangleq a$ for any $a\in A_x$ such that $z_{xa}^*>0$. If $z_{xa}^*=0$ for all $a\in A_x$, choose an arbitrary $a\in A_x$, and set $\tilde{u}^*(x)=a$.

$\tilde{u}^*$ With Constraints on Data and Voice Blocking: Let $c(x,a)$, which is defined in (28), be some suitably defined cost function that reflects the rate at which the BS incurs administrative costs when it chooses action a while in state x. Minimizing the continuous average cost (29) subject to constraints on the maximum allowable data and voice blocking probability (for each voice class) is easily addressed by adding the following (sample path) constraints to the LP in (39):

$$\sum_{x\in X}\sum_{a\in A_x}c_d(x,a)\tau_x(a)z_{xa}\leq\gamma_d \tag{40}$$

$$\sum_{x\in X}\sum_{a\in A_x}c^i(x,a)\tau_x(a)z_{xa}\leq\gamma_v^i,\quad i=1,\ldots,K_v \tag{41}$$

where c_d is defined in Proposition III.1, γ_d is the maximum allowed data-blocking probability, $c^i(x,a)=1-a_{v,i}$ [see (33)], and γ_v^i is the maximum allowed blocking probability for voice class i. [This is a nice feature of the LP approach for solving the CAC admission problem (29) and (30) that is not available with the policy iteration and value iteration methods.] Let z_{xa}^* denote the optimal solution to the LP. When sample path constraints are included, the optimal policy will in general be a *randomised stationary policy*: The optimal action for state x is the control a chosen from the set A_x probabilistically according to the probabilities $\tau_x(a)z_{xa}/\sum_{a\in A_x}\tau_x(a)z_{xa}$, $a\in A_x$. If $z_{xa}^*=0$ for all $a\in A_x$, choose an arbitrary $a\in A_x$, and set $\tilde{u}^*(x)=a$. See [19] for details.

[4]See [1, Ch. 5] and [19, Ch. 3] for a discussion on the value iteration and policy iteration approaches. In [17], the computational savings are examined when the *modified value iteration* algorithm is used instead of the LP approach.

V. Implementation Issues and Extensions

Constructing the State Space: As described in Section III, the SIR constraints define the state space of the SMDP (17). Due to Proposition IV.1, the procedure for constructing X is simple and is as follows.

Step 1) Construct an upper bound $\bar{X}$ for X, i.e., $X \subset \bar{X}$ as follows: Let $M_d \triangleq \max\{x_d \in \mathbb{Z}_+ : \Psi_d(x_d e_1) \geq \beta_d\}$, $M_v^i \triangleq \max\{x_{v,i} \in \mathbb{Z}_+ : \Psi_{v,i}(x_{v,i} e_{i+2}) \geq \beta_{v,i}\}$, $i = 1, \ldots, K_v$. [e_i is defined in (21).] Note that $\Psi_d(x_d e_1)$ and $\Psi_{v,i}(x_{v,i} e_{i+2})$ are monotonically decreasing in x_d and $x_{v,i}$ respectively; this follows by Corollary A.2. M_d is the maximum number of active data users that can be supported simultaneously, whereas M_v^i is the maximum number of active class i users that can be supported. Let $\bar{X} \triangleq \{x = [x_d, x_b, x_{v,1}, x_{v,2}, \ldots, x_{v,K_v}]^T \in \mathbb{Z}_+{}^{2+K_v} : x_d \leq M_d,\ x_b \leq B_d,\ x_{v,i} \leq M_v^i, i = 1, \ldots, K_v\}$. Using Proposition IV.1, one may verify that $X \subset \bar{X}$.

Step 2) "Trim" $\bar{X}$ down to X as follows: For each $x \in \bar{X}$, if $\Psi_d(x) \geq \beta_d$ and $\Psi_{v,i}(x) \geq \beta_{v,i}$, $i = 1, \ldots, K_v$, then $x \in X$. By Proposition IV.1, if $x \in X$, then all $x' \in \bar{X}$ such that $x' \leq x$ are also in X. Similarly, if $x \notin X$, then all $x' \in \bar{X}$ such that $x' \geq x$ are also not in X. Step 2 involves a finite number of evaluations that is bounded above by $|\bar{X}|$.

Computational Complexity: To construct the optimal CAC policy $\tilde{u}^*$(30), one needs to first i) construct the state space X(17) and then ii) solve the LP in (39). The task of constructing the optimal CAC policy is entirely offline as i) and ii) are offline procedures. The computational complexity of constructing $\tilde{u}^*$ is the computational complexity of procedures i) and ii). As indicated in Step 2, constructing X involves a finite number of *SIR feasibility evaluations* that are bounded above by $|\bar{X}|$; note that this is the worst-case scenario. Although one needs to verify that each point in $\bar{X}$ also belongs to X, the convexity property in Proposition IV.1 greatly simplifies this procedure, as detailed in Step 2. The cardinality of $\bar{X}$ itself depends on the dimension of the state vector $x(t)$(15), the size of the data buffer B_d, and the particular value of constants $M_d,\ M_v^i, i = 1, \ldots, K_v$ given in Step 1. The complexity of procedure ii) is polynomial in the number of decision variables of the LP (39), and the number of decision variable in the LP are bounded above by $|X||A|$. (Note that an LP may be solved in polynomial time by so-called *interior point methods*.) The main approach to reducing the complexity of procedures i) and ii) would be to reduce the cardinality of $\bar{X}$ and the state space X, respectively. $|\bar{X}|$ and $|X|$ are simultaneously reduced by reducing the dimensions of the state vector $x(t)$(15). This, in turn, can be achieved by reducing number of user classes K_v under consideration.

Minimizing the Dropping Probability of Handoff Calls: As in [5], we can construct an optimal CAC policy that minimizes the probability of dropping handoff calls or including handoff calls dropping probability as a constraint. The basic idea is a follows: As in [5], we assume handoff calls from adjacent cells arrive according to Poisson processes with rates $\lambda_{d,hoff}$ and $\lambda_{v,i,hoff}$, $i = 1, \ldots, K_v$. Note that the total rate of incoming calls class i is now $\lambda_{v,i} + \lambda_{v,i,hoff}$, i.e., the sum of the rates of new and handoff calls. The definition of the state space remains the same as in (17). This is because we do not distinguish between an active new class i and an active handoff class i as the same bandwidth (SIR requirement) is consumed. To explicitly minimize the handoff call blocking probability, we need to redefine the action space in (19) as follows:

$$A = \{a = (a_d, a_{d,hoff}, a_b, a_{v,1}, a_{v,1,hoff} \ldots, a_{v,K_v}, a_{v,K_v,hoff}) : a \in \{0,1\}^{3+2K_v}\}. \quad (42)$$

$a_{v,i,hoff} = 1$ implies that we accept an arriving handoff class i call; otherwise, we reject it. (There is a similar interpretation for an arriving handoff data call. There is no buffering for handoff data calls.) Because the action space has been redefined, the set of feasible state-action pairs in (21), the transition probabilities in (26), and the mean holding time in (25) will have to be changed accordingly; compare with [5], (8) and (15). To minimize the handoff call blocking probability, we set $c(x,a) = 1 - a_{v,i,hoff}$ in (29).

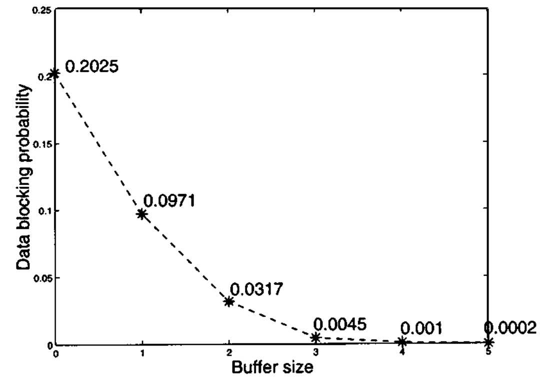


Fig. 2. Data-blocking probability versus buffer size—Values indicated adjacent.

Multiple Data Classes: We may generalize the CAC problem to the case when the transmit power, channel gain, and variance of data users are quantized to one of a fixed set of K_d different values: $\{(P_{d,i}, \bar{h}_{d,i}, \xi_{d,i}^2), \quad i = 1, \ldots, K_d\}$. The basic idea to performing optimal CAC for this scenario is illustrated in Fig. 2. An arriving data user is admitted into a common buffer, provided that there is space. (Note that all data users are buffered by default.) The state of the BS in (15) is now redefined as follows:

$$x(t) = [x_{d,1}(t), \ldots, x_{d,K_d}(t), x_{b,1}(t), \ldots, x_{b,K_d}(t), x_{v,1}(t), \ldots, x_{v,K_v}(t)]^T$$

where $x_{d,i}$ and $x_{v,i}$ denote the number of active class i data and voice users, respectively, and $x_{b,i}$ denotes the number of class i data users in the buffer. The CAC problem may then be reformulated to admit queued data users, as was done in Section III.

VI. Numerical Examples

In the section, the performance of the optimal CAC policy is studied by addressing the following issues: i) Is there a notable reduction in the data-blocking probability if a nonzero buffer size is employed? ii) For a finite data buffer size, does the

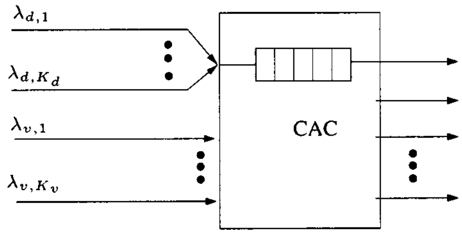


Fig. 3. Call admission for multiple voice and data classes.

policy that minimizes the data-blocking probability also minimize the average waiting time in the buffer? The problem of minimizing the data-blocking probability subject to constraints on the maximum voice blocking probability for each class is also considered.

Simulation Parameters: In all the simulations that follow, the parameters used in the integrated voice/data DS-CDMA system model are as follows: two classes of voice users [$K_v = 2$ in (1)] with transmit powers $\{P_{v,1}, P_{v,2}\} = \{1.2, 1.7\}$ and minimum SIR thresholds (16) $\{\beta_{v,1}, \beta_{v,2}\} = \{110{,}150\}$. The transmit power and SIR threshold of data users are $P_d = 2.5$ and $\beta_d = 200$, respectively. The data SIR threshold is larger than the voice thresholds since voice is more tolerant to transmission errors. Class 2 voice users are *premium* users, whereas class 1 users are *basic*. The arrival and departure rates of the voice and data users are as follows (3, 4): $\lambda_d = 1$, $\{\lambda_{v,1}, \lambda_{v,2}\} = \{0.5, 0.5\}$, $\mu_d = 2$, $\{\mu_{v,1}, \mu_{v,2}\} = \{1.1, 1.1\}$. The buffered data admit latency rate $\lambda_b = 200(24)$. The parameters of the signal model in (5) are as follows: $N = 32$, $\sigma^2 = 0.01$. In (1) and (2), $\bar{h}_d = \bar{h}_{v,1} = \bar{h}_{v,2} = 1$ and $\xi_d^2 = \xi_{v,1}^2 = \xi_{v,2}^2 = 0$, i.e., the channel is assumed to be known perfectly.

A. Data-Blocking Probability versus Buffer Size

This section considers the effect of buffer length on the data-blocking probability. The CAC policy that minimized the data-blocking probability subject to the following values for the maximum voice-blocking probability was constructed for various values of the data buffer size B_d (17): class 1 blocking probability ≤ 0.1, class 2 blocking probability ≤ 0.1. In effect, the LP in (39) was solved with c given in Proposition III.1 and sample path constraints (41). The minimum value for the data-blocking probability as B_d is varied is presented in Fig. 3. The results point to the improvements yielded with buffering. Note the geometric reduction in blocking probability; this implies that beyond a particular size for the buffer, the reduction in data-blocking probability that can be expected is marginal.

B. Waiting Time in the Data Buffer

This section considers waiting time in the queue. It is not possible to write the waiting time in the data buffer as an additive cost function, as was done for the data-blocking probability in Proposition III.1. Therefore, the LP in (39) cannot be used to construct a policy that minimizes the average waiting time in the data buffer. However, since the data buffer is finite, it seems reasonable to expect that the CAC policy that minimizes the average waiting time in the data buffer would correspond to the policy that minimizes the data-blocking probability. We demonstrate this to be untrue in the simulation that follows.

TABLE I
INTERPLAY BETWEEN DATA-BLOCKING PROBABILITY AND WAITING TIME

$[\nu_1, \nu_2, \nu_3]$	Data blocking probability	Average waiting time in data buffer (s)
[.2 , .4, .4]	0.0629	1.5748
[.4 , .3, .3]	0.0490	1.2289
[.6 , .2, .2]	0.0335	1.0952
[.8 , .1, .1]	0.0176	0.8846
[.9 , .05, .05]	0.0092	0.8167
[1 , 0, 0]	8e-08	1.1892

The CAC policy that minimized a weighted sum of the data and voice-blocking probability was constructed for various values of the weighting factors ν_1, ν_2, and ν_3 by solving LP (39) with c defined in (34). Table I records the values used for the weighting factors as well as the corresponding data-blocking probability and average waiting time in the data buffer. The average waiting time was obtained by simulating a BS where data and voice users arrived, sojourned, and departed according to their appropriate processes. The CAC policy obtained by solving the LP was used to control admission. The average waiting time was obtained by averaging the waiting time of all data users that were admitted into the data buffer.

The solution to the LP when $\nu_1 = 1$ corresponds to the CAC policy that yields the minimum data-blocking probability. Note that the average waiting time decreases as nu_1 increases to 0.9 and then increases. The policy corresponding to $\nu_1 = 1$, although yielding minimum data-blocking probability, does not minimize the waiting time in the data buffer.

C. Data-Blocking Probability versus Voice-Blocking Probability

In this section, the CAC policy that minimizes that data-blocking probability subject to constraints on the voice blocking probability is constructed, i.e., LP (39) is solved with c given in Proposition III.1 and sample path constraints (41). The data buffer size is held at 5. Table II records the minimum data-blocking probability that can be had for various values of $[\gamma_v^1, \gamma_v^2]$. As the maximum allowed voice blocking probability for both the premium and standard class users are decreased, as expected, the data-blocking probability increases.

TABLE II
INTERPLAY BETWEEN DATA-BLOCKING PROBABILITY AND VOICE-BLOCKING CONSTRAINTS. BLOCKING PROBABILITY OF PREMIUM USERS IS HELD AT 0.05 BELOW STANDARD USERS

Voice blocking probability constraints $[\gamma_v^1, \gamma_v^2]$	Data blocking probability
[0.17 0.12]	4.2e-5
[0.15 0.1]	0.0001
[0.13 0.08]	0.0001
[0.11 0.06]	0.0097
[0.09 0.04]	0.8887

Note, however, as indicated in the results of Section VI-A, that the data-blocking probability can be further decreased by increasing the buffer size.

VII. CONCLUSIONS

Within the SMDP framework, the CAC problem was solved for a multiservice DS-CDMA cellular system transmitting in a fading environment with LMMSE receivers at the BS. Blocking probabilities (data and voice) were shown to be represented by additive cost functions. The optimal CAC policy was constructed using an LP, and blocking-probability constraints for the various user classes were accommodated by merely adding additional linear constraints to this LP. A buffering scheme was proposed to cope with congestion. In numerical examples, the gains, in terms of the reduction of the data-blocking probability, were demonstrated to be significant, even for small buffer sizes. Suboptimal algorithms (for complexity reduction) for solving SMDPs based on state aggregration is a subject for future research.

APPENDIX
PROOFS

Proof: [Proposition II.1]: Using (9), write the constraint on $\widehat{\text{SIR}}_1$ as a constraint on β as follows:

$$\beta \geq \frac{\underline{\text{SIR}}_1}{P_1|\bar{h}_1|^2 - \underline{\text{SIR}}_1 P_1 \xi_1^2}. \tag{43}$$

(Assume the denominator is positive or else the constraint $\widehat{\text{SIR}}_1 \geq \underline{\text{SIR}}_1$ cannot be satisfied since, from (9), $\widehat{\text{SIR}}_1 \leq |\bar{h}_1|^2/\xi_1^2$.) Now, to verify that the fixed point β in (9) satisfies the constraint in (43), use the following necessary and sufficient condition that is given in Lemma A.1: For any $\gamma > 0$

$$\beta \geq \gamma \iff [\sigma^2 + \frac{1}{N}\sum_{k=2}^{K} I\left(P_k(\xi_k^2 + |\bar{h}_k|^2), \gamma\right)]^{-1} \geq \gamma. \tag{44}$$

■

The following proposition is needed in the proof of Proposition III.1. The limit in (29) is independent of x_0 when $\tilde{u}$ is a *unichain* policy. A policy $\tilde{u}$ is said to be unichain if the corresponding *embedded Markov chain* has no two disjoint closed sets of states. The embedded Markov chain corresponding to $\tilde{u}$ is the controlled discrete time Markov chain $\{x(t_k)\}_{k\in\mathbb{Z}_+}$ (t_k defined in (18)) with transition probabilities $p_{xy}(\tilde{u}(x))$ (22). A nonempty set $\tilde{X} \subseteq X$ is said to be closed if $p_{xy}(\tilde{u}(x)) = 0$ when $x \in \tilde{X}$ and $y \notin \tilde{X}$; see also [1], [19].

Proposition A.1: All $\tilde{u} \in \mathcal{U}$ are unichain.

Proof: Consider any $\tilde{u} \in \mathcal{U}$. We write $x \to y$ if state y is accessible from x, i.e., there exists a sequence of states $x_1, x_2, \ldots, x_k \in X$ for some finite k such that $p_{xx_1}(\tilde{u}(x))p_{x_1x_2}(\tilde{u}(x_1))\ldots p_{x_{k-1}x_k}(\tilde{u}(x_{k-1}))p_{x_ky}(\tilde{u}(x_k)) > 0$.

Case 1) $B_d = 0$. Consider any $x \in X$ such that $x \neq [0,0,\ldots,0]^T$. Since $p_{x(x-e_i)}(\tilde{u}(x)) > 0$ for all $i \in \{1,3,\ldots,2+K_v\}$ such that $x - e_i \in X$, we have $x \to [0,0,\ldots,0]^T$. Thus, state $[0,0,\ldots,0]^T$ is accessible from all states in X, and therefore, no two disjoint closed set of states exist.

Case 2) $B_d > 0$. For $i \in \{0,1,\ldots,B_d\}$, let $X_i = \{x \in X : x_b = i\}$, $x_i = [0,i,0,\ldots,0]^T$. Clearly, $X = \bigcup_{i=0}^{B_d} X_i$. We will show that state x_{B_d} is accessible from all states in X and, therefore, that no two disjoint closed set of states can exist. As in Case 1, it may be shown that $x \to x_i$ for all $x \in X_i \backslash x_i$, $i \in \{0,\ldots,B_d\}$. In addition, $x_i \to x_{i+1}$, $i \in \{1,\ldots,B_d-1\}$. To see why, from (22), $p_{x_ix_{i+1}}(\tilde{u}(x_i)) = \lambda_d(1-a_d)\delta(B_d - i)\tau_{x_i}(\tilde{u}(x_i))$. Since $i > 0$, $(1-a_d) = 1$, by virtue of (21). $\delta(B_d - i) = 1$ since $i < B_d$. Thus, $\tau_{x_i}(\tilde{u}(x_i)) > 0$ (25), and therefore, $p_{x_ix_{i+1}}(\tilde{u}(x_i)) > 0$. Finally, one must show that there exists some $x \in X_1$ such that $x_0 \to x$. [If for some $x \in X_0$ and $y \notin X_0$, $p_{xy}\tilde{u}(x)) > 0$, obviously, $y \in X_1$ as transitions are only possible to neighboring states (26)] Let $\tilde{X}_0 = \{x \in X_0 \backslash x_0 : x_0 \to x\}$. If $\tilde{X}_0 = \emptyset$, then $p_{x_0x_1}(\tilde{u}(x_0)) = 1$, i.e., $x_0 \to x_1$. Therefore, assume $\tilde{X}_0 \neq \emptyset$. If there exists $x \in \tilde{X}_0$ such that $\tilde{u}(x)$ satisfies $a_d = 0$, then, from (26), $p_{x(x+e_2)}(\tilde{u}(x)) > 0$. If for all $x \in \tilde{X}_0$, $\tilde{u}(x)$ satisfies $a_d = 1$, from (26), the state $y = [x_d, 0, 0, \ldots, 0]^T \in \tilde{X}_0$ such that $y + e_1 \notin X$ (y corresponds to the state that has the maximum number of permissible active data users) is accessible from x_0. From A_y (21) and (26), we note that $p_{y(y+e_2)}(\tilde{u}(y)) > 0$. Thus, $x_0 \to y + e_2$. ■

Proof: [Proposition III.1]: The proof that follows is for Case i). Cases ii) and iii) follow similar arguments.

The sample paths of the process $\{x(t)\}_{t\in\mathbb{R}_+}$ are *continuous on the right and have limits on the left* (corlol) w.p. 1. This follows because of the following.

a) New data and voice users arrive according to homogeneous Poisson processes (3), which is a *nonexplosive point process* (see [2] for a definition of a nonexplosive point process).

b) We have adopted the convention that the occurrence of a *trigerring event* at some time s will cause a state transition in the process $\{x(t)\}_{t\in\mathbb{R}_+}$ at time s itself; triggering events here are the arrival of a data or voice user that is admitted, a data user in the buffer is made active at the service completion of an active user.

Since $a(t) = \tilde{u}(x(t))$ (29), it follows that sample paths of $\{a(t)\}_{t\in\mathbb{R}_+}$ are also corlol w.p. 1. For $s > 0$, let $a(s-)$ and $x(s-)$ denote the left-hand limit of $a(t)$ and $x(t)$ at s, respectively, i.e., $a(s-) = \lim_{h\to s, h<s} a(h)$. Set $a(0-) = a(0)$ and $x(0-) = x(0)$. Let $\{t_k^d\}_{k\in\mathbb{Z}_+}$ denote the arrival instance of the data users with $t_0^d = 0$ by convention. Note that by (31)

$$\bar{N}_d(T) = \sum_{k=1}^{N_d(T)} c(x(t_k^d-), a(t_k^d-)) \tag{45}$$

since an arriving data user is blocked if and only if $a(t_k^d-) \in A$ is such that $a_d = 0$ and the data buffer is full, i.e., $x(t_k^d-) + e_2 \notin X$. Note that samples paths of $\{c(x(t-), a(t-))\}_{t\in\mathbb{R}_+}$ are left continuous and have limits on the right w.p. 1. Since $\{N_d(t)\}_{t\in\mathbb{R}_+}$ is a Poisson process (3), $\{N_d(t+s) - N_d(t) : s \geq 0\}$ is independent of $\{x(s) : 0 \leq s \leq t\}$ for all $t \geq 0$. This is true due to the *memoryless* property of Poisson processes. Thus, $\{N_d(t+s) - N_d(t) : s \geq 0\}$ is also independent of $\{c(x(s-), a(s-)) : 0 \leq s \leq t\}$ as a and c are measurable functions of x; this property is called the *lack of anticipation assumption* (LAA); see [7, Def. 7.2]. (Note that the sigma algebra $\sigma\{c(x(s-), a(s-)) : 0 \le s \le t\} \subset \sigma\{c(x(s), a(s)) : 0 \le s \le t\} \subset \sigma\{x(s) : 0 \le s \le t\}$.) Because LAA holds, by [7, Th. 7.3], the well-known *Poisson arrivals see time averages* property holds, which asserts that

$$\lim_{T\to\infty} \frac{\sum_{k=1}^{N_d(T)} c(x(t_k^d-), a(t_k^d-))}{N_d(T)} = \lim_{T\to\infty} \frac{1}{T}\int_0^T c(x(t), a(t))\, dt \quad \text{w.p. } 1 \tag{46}$$

provided either limit exists, i.e., existence of one limit implies existence of the other, and both limits are the same. The proof is complete by noting the following properties of unichain policies [19, Th. 3.5.1]: First, $\{x(t_k)\}_{k\in\mathbb{Z}_+}$ has a unique stationary distribution $\{\pi_{\tilde{u}}(x), x \in X\}$. Second, for each initial state x_0, $\lim_{T\to\infty} 1/T \int_0^T c(x(t), a(t))dt = J_{\tilde{u}}$ w.p. 1, where

$$J_{\tilde{u}} = \frac{\sum_{x\in X} c(x, \tilde{u}(x))\tau_x(\tilde{u}(x))\pi_{\tilde{u}}(x)}{\sum_{x\in X} \tau_x(\tilde{u}(x))\pi_{\tilde{u}}(x)}. \tag{47}$$

Third, $J_{\tilde{u}}(x_0) = J_{\tilde{u}}$ for all x_0. ■

Lemma A.1: Consider $p_1, p_2, \ldots, p_k > 0$, $N > 0$, $\sigma^2 \ge 0$. There exists a unique positive β that satisfies

$$\beta = \left[\sigma^2 + \frac{1}{N}\sum_{i=1}^{k} I(p_i, \beta)\right]^{-1} \tag{48}$$

if $\sigma^2 > 0$ or $(k-1)/N > 1$; I is defined in (9). Furthermore, for any $\gamma > 0$

$$\beta > \gamma \iff \left[\sigma^2 + \frac{1}{N}\sum_{i=1}^{k} I(p_i, \gamma)\right]^{-1} > \gamma. \tag{49}$$

A consequence of this lemma is the following.

Corollary A.2: Consider $p_1, p_2, \ldots, p_k > 0$, $k > 1$, $N > 0$, $\sigma^2 \ge 0$. Let β_{k-1} and β_k be the unique positive fixed points that satisfy

$$\beta_{k-1} = \left[\sigma^2 + \frac{1}{N}\sum_{i=1}^{k-1} I(p_i, \beta_{k-1})\right]^{-1}$$
$$\beta_k = \left[\sigma^2 + \frac{1}{N}\sum_{i=1}^{k} I(p_i, \beta_k)\right]^{-1}. \tag{50}$$

Then, $\beta_{k-1} > \beta_k$.

Proof: From Lemma A.1, $\beta_{k-1} > \beta_k$ iff $[\,\sigma^2 + 1/N \sum_{i=1}^{k-1} I(p_i, \beta_k)\,]^{-1} > \beta_k$. The result follows since β_k satisfies (50), and $I(p_k, \beta_k) > 0$. ■

Proof of [Proposition IV.1] Part i): The result is obvious when $e_j = e_2$. In the following, we consider any $e_j \neq e_2$. To show $x \in X$, by the definition of X (17), we must show

$$\Psi_d(x) \ge \beta_d, \quad \Psi_{v,i}(x) \ge \beta_{v,i}, \quad i = 1, \ldots, K_v. \tag{51}$$

We establish (51) for $\Psi_d(x) \ge \beta_d$; the same idea may be repeated to show $\Psi_{v,i}(x) \ge \beta_{v,i}$.

Case 1) x is such that $x_d = 0$. $\Psi_d(x) \ge \beta_d$ by definition.

Case 2) $x = e_1 + me_2$ where m is any non-negative integer such that $m \le B_d$. By hypothesis and the definition of X, $\Psi_d(e_1 + me_2 + e_j) \ge \beta_d$. It suffices to show that $\Psi_d(e_1 + me_2) \ge \Psi_d(e_1 + me_2 + e_j)$. Let $\beta > 0$ satisfy $\beta = f_d(e_1 + me_2 + e_j, \beta)$. Obviously, $1/\sigma^2 > \beta$, and therefore, $\Psi_d(e_1 + me_2) > \Psi_d(e_1 + me_2 + e_j)$.

Case 3) This involves any x not covered by Cases 1 and 2. It suffices to show $\Psi_d(x) \ge \Psi_d(x + e_j)$. Let β and β' satisfy $\beta = f_d(x, \beta)$, $\beta' = f_d(x + e_j, \beta')$. By Corollary A.2, $\beta > \beta'$, and therefore, $\Psi_d(x) > \Psi_d(x + e_j) \ge \beta_d$.

Part ii): The proof follows similar arguments to those used in Part i). ■

REFERENCES

[1] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Belmont, CA: Athena Scientific, 1995, vol. 2.

[2] P. Bremaud, *Markov Chains, Gibbs Fields, Monte Carlo Simulation, and Queues*. New York: Springer-Verlag, 1999.

[3] C. G. Cassandras and S. Lafortune, *Introduction to Discrete Event Systems*. Boston, MA: Kluwer, 1999.

[4] I. Chih-Lin and R. D. Gitlin, "Multi-code CDMA wireless personal communications networks," in *Proc. IEEE ICC*, 1995, pp. 1060–1064.

[5] J. Choi, T. Kwon, Y. Choi, and M. Naghshineh, "Call Admission control for multimedia services in mobile cellular networks: a Markov decision approach," in *Proc. IEEE Int. Symp. Comput. Commun.*, Antibes, Greece, July 2000.

[6] C. Comaniciu, N. B. Mandayam, D. Famolari, and P. Agrawal, "QoS guarantees for third generation (3G) CDMA systems via admission and flow control," in *Proc. IEEE Veh. Technol. Conf.*, Boston, MA, Sep. 2000.

[7] J. H. Dshalalow, *Advances in Queueing: Theory, Methods and Open Problems*. Boca Raton, FL: CRC, 1995.

[8] J. Evans and D. N. C. Tse, "Large system performance of linear multiuser receivers in multipath fading channels," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2059–2078, Sept. 2000.

[9] M. Honig and J.-B. Kim, "Outage probability of multi-code DS-CDMA with linear interference suppression," in *Proc. IEEE Military Commun. Conf.*, Bedford, MA, Oct. 1998, pp. 248–252.

[10] M. L. Honig and H. V. Poor, "Adaptive interference suppression," in *Wireless Communications: Signal Processing Perspectives*, H. V. Poor and G. W. Wornell, Eds. Upper Saddle River, NJ: Prentice-Hall, 1998.

[11] T. H. Lee and J. T. Wang, "Admission control for VSG-CDMA systems supporting integrated services," in *IEEE Proc. GLOBECOM*, Sydney, Australia, Nov., pp. 2050–2055.

[12] T. Liu and J. A. Silvester, "Joint admission/congestion control for wireless CDMA systems supporting integrated services," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 845–857, Aug. 1998.

[13] Z. Liu and M. E. Zarki, "SIR-based call admission control for DS-CDMA cellular systems," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 638–644, May 1994.

[14] S. L. Miller, M. L. Honig, and L. B. Milstein, "Performance analysis of MMSE receivers for DS-CDMA in frequency-selective fading channels," *IEEE Trans. Commun.*, vol. 48, pp. 1919–1929, Nov. 2000.

[15] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York: Wiley, 1994.

[16] K. W. Ross, *Multiservice Loss Models for Broadband Telecommunication Networks*. London, U.K.: Springer-Verlag, 1995.

[17] K. W. Ross and D. Tsang, "Optimal circuit access policies in an ISDN environment: a Markov decision approach," *IEEE Trans. Commun.*, vol. 37, no. 9, pp. 934–939, Sept. 1989.

[18] M. Soroushnejad and E. Geraniotis, "Multi-access strategies for an integrated voice/data CDMA packet radio network," *IEEE Trans. Commun.*, vol. 43, pp. 934–945, Feb. 1995.

[19] H. C. Tijms, *Stochastic Models: An Algorithmic Approach*. Chichester, U.K.: Wiley, 1994.

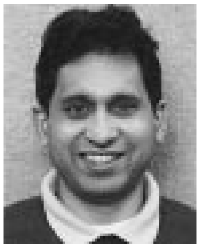
[20] D. Tse and S. V. Hanly, "Linear multiuser receivers: effective interference, effective bandwidth and user capacity," *IEEE Trans. Inform. Theory*, vol. 45, pp. 641–657, Mar. 1999.

[21] W. Yang and E. Geraniotis, "Admission policies for integrated voice and data traffic in CDMA packet radio networks," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 654–664, May 1994.

[22] J. Zhang, E. K. P. Chong, and D. Tse, "Output MAI distributions of linear MMSE multiuser receivers in CDMA systems," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1128–1144, Mar. 2001.

Sumeetpal Singh is currently pursuing the Ph.D. degree at the University of Melbourne, Parkville, Australia.

His research interests are in stochastic systems and statistical signal processing.



Vikram Krishnamurthy (SM'99) was born in 1966. He received the B.E.E. degree from the University of Auckland, Auckland, New Zealand, in 1988 and the Ph.D. degree from the Australian National University, Canberra, in 1992.

He is currently a Professor with the Department of Electrical Engineering, University of Melbourne, Parkville, Australia, and serves as Deputy Head of the department. Starting in August 2002, he will be a Professor with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, Canada. His research interests span several areas including stochastic scheduling and network optimization, time-series analysis, and statistical signal processing. He is an Associate Editor of *Systems and Control Letters*.

Dr. Krishnamurthy is currently an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He has served on the technical program committee of several conferences including the 37th IEEE Conference on Decision and Control, held in Tampa, FL, in 1998, and the IFAC Symposium on System Identification in 2000 (SYSID'2000), held in Santa Barbara, CA.

H. Vincent Poor (S'72–M'77–SM'82–F'87) received the Ph.D. degree in electrical engineering and computer science in 1977 from Princeton University, Princeton, NJ.

He is currently a Professor of electrical engineering at Princeton. He is also affiliated with Princeton's Department of Operations Research and Financial Engineering and with its Program in Applied and Computational Mathematics. From 1977 until he joined the Princeton faculty in 1990, he was a Faculty Member at the University of Illinois at Urbana-Champaign. He has also held visiting and summer appointments at several universities and research organizations in the United States, Britain, and Australia. His research interests are primarily in the area of statistical signal processing, with applications in wireless communications and related areas. Among his publications in this area is the book *Wireless Communications: Signal Processing Perspectives*.

Dr. Poor is a member of the National Academy of Engineering and is a Fellow of the Acoustical Society of America, the American Association for the Advancement of Science, the Institute of Mathematical Statistics, and the Optical Society of America. His IEEE activities include serving as the President of the IEEE Information Theory Society in 1990 and as a member of the IEEE Board of Directors from 1991 to 1992. Among his honors are the Frederick E. Terman Award of the American Society for Engineering Education (1992), the Distinguished Member Award from the IEEE Control Systems Society (1994), an IEEE Third Millennium Medal (2000), the IEEE Graduate Teaching Award (2001), and the first Joint Paper Award of the IEEE Communications Society and IEEE Information Theory Society (2002).