Constrained Maximum-Likelihood Detection in CDMA

Peng Hui Tan, Student Member, IEEE, Lars K. Rasmussen, Member, IEEE, and Teng J. Lim, Member, IEEE

Abstract—The detection strategy usually denoted optimal multiuser detection is equivalent to the solution of a (0, 1)-constrained maximum-likelihood (ML) problem, a problem which is known to be NP-hard. In contrast, the unconstrained ML problem can be solved quite easily and is known as the decorrelating detector. In this paper, we consider the constrained ML problem where the solution vector is restricted to lie within a closed convex set (CCS). Such a design criterion leads to detector structures which are ML under the constraint assumption. A close relationship between a sphere-constrained ML detector and the well-known minimum mean square error detector is found and verified. An iterative algorithm for solving a CCS constraint problem is derived based on results in linear variational inequality theory. Special cases of this algorithm, subject to a box-constraint, are found to correspond to known, nonlinear successive and parallel interference cancellation structures, using a clipped soft decision for making tentative decisions, while a weighted linear parallel interference canceler with signal-dependent weights arises from the sphere constraint. Convergence issues are investigated and an efficient implementation is suggested. The bit-error rate performance is studied via computer simulations and the expected performance improvements over unconstrained ML are verified.

Index Terms—Code-division multiple access, interference cancellation, multiuser detection.

I. INTRODUCTION

In any multiple-access system, the available resources are shared in some way among all active users. As a consequence, there is a fundamental tradeoff between the amount of resources available for each user and the corresponding interference encountered due to multiple access. In code-division multiple-access (CDMA) systems all resources are in principle available to all users simultaneously. The users are distinguished from each other by user-specific signature sequences, modulating the transmitted data symbols using direct-sequence spread-spectrum techniques. This in turn leads to a relative high level of multiple-access interference (MAI) as it is not feasible to maintain low (or zero) cross correlation among all users in a practical random-access system.

Conventional spread-spectrum detection techniques applied in CDMA are severely limited in performance by MAI, leading to both system capacity limitations as well as strict power control requirements [1]. These limitations are due to the fact that the traditional matched filter output does not represent a sufficient statistic for detection. A detector working on a true sufficient statistic is generally denoted a multiuser detector, and it has the potential of alleviating the MAI problems encountered by conventional techniques.

In order to describe the detection strategies to follow, assume an asynchronous transmission of \( L \) information bits per user using binary phase-shift keying (BPSK) modulation. The number of active users is \( K \) and the data vector consisting of all transmitted data symbols for all users is denoted by the column vector \( \mathbf{d} \) of dimension \( LK \). The general maximum-likelihood (ML) detection problem is equivalent to a constrained quadratic optimization. The maximally constrained ML detector finds the ML solution constrained to \( \mathbf{d} \in \{-1, 1\}^{LK} \) where \( \{-1, 1\}^{LK} \) denotes the set of all binary \( LK \)-tuples represented as column vectors, i.e., each information symbol estimate must be either +1 or −1. This detector has previously been denoted the optimal multiuser detector [2]. In the area of optimization, the above ML problem is known as a (0, 1)-constrained (or Boolean-constrained) quadratic minimization which in turn represents a combinatorial quadratic minimization. Such a problem is known to be NP-hard [3] so the (0, 1)-constrained ML detector is therefore in general too complex for practical asynchronous DS-CDMA systems, even with a moderate number of users. For certain special cases of the correlation matrix\(^1\) it has been shown that (0, 1)-constrained ML detection can be obtained by successive interference cancellation [4] or by polynomial-time algorithms [5]–[7]. A class of complexity-limiting (0, 1)-constrained ML detectors was suggested in [8], assuming a tree-search based detector structure. An iterative structure which is guaranteed to deliver (0, 1)-constrained ML decisions on some bits was suggested in [9]. The matched filter outputs are here compared to an iteratively tightened threshold through which (0, 1)-constrained ML decisions are made. Decisions on all bits are however not guaranteed. Approximations to the (0, 1)-constrained ML problem have also

\(^1\)The case of all identical cross correlations, i.e., identical off-diagonal elements of the correlation matrix.

Paper approved by G. Caire, the Editor for Multiuser Detection and CDMA of the IEEE Communications Society. Manuscript received September 1, 1999; revised January 6, 2000 and June 9, 2000. This work was supported in part by the Centre for Wireless Communications (Singapore), the National University of Singapore, the Swedish Research Council for Engineering Sciences (TFR), and the Swedish Foundation for Strategic Research (SSF). This paper was presented in part at the International Conference on Information, Communication and Signal Processing, Singapore, December 1999, the 2000 International Zurich Seminar on Broadband Communication, Zurich, Switzerland, February 2000, the IEEE Vehicle Technology Conference, Tokyo, Japan, May 2000, and the IEEE International Symposium on Information Theory, Sorrento, Italy, June 2000.

P. H. Tan and L. K. Rasmussen are with the Telecommunication Theory Group, Department of Computer Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden (e-mail: phtan@ce.chalmers.se; larsr@ce.chalmers.se).

T. J. Lim is with the Centre for Wireless Communications, Singapore Science Park II, Singapore 117674 (e-mail: cwclim@nus.edu.sg).

Publisher Item Identifier S 0090-6778(01)00257-4.
been suggested in [10] based on the expectation maximization algorithm and in [11] based on iterative transformations of the quadratic minimization problem such that the unconstrained solution to the transformed problem monotonically approaches the desired solution. In this paper, however, we will take a more general approach to complexity-limiting ML detection.

To reduce complexity, the constraints imposed on a feasible solution can be relaxed. A simple constraint to impose is to restrict the solution vector to be contained within a closed convex set (CCS). Examples of CCSs of dimension \( LK \) are \( \mathbb{R}^{LK} \), an ellipsoid of dimension \( LK \) and a hypercube of dimension \( LK \). The corresponding optimization problem is known as a CCS constrained quadratic program (CCSQP). The fully unconstrained ML detector was suggested in [12] and is denoted the decorrelating detector. Here, a valid solution vector \( \mathbf{u} \) is found in \( \mathbb{R}^{LK} \) as each symbol estimate can take on any real value, i.e., no constraints are imposed. The case is denoted an unconstrained quadratic program (UQP). This ML solution is then mapped onto a valid data point through \( \hat{\mathbf{d}} = \text{Sgn}(\mathbf{u}) \). This is of course a suboptimal mapping and \( \hat{\mathbf{d}} \) is not ML [13].

Since the length of the data symbol vector \( \mathbf{d} \) representing the transmitted symbols is constant for constant envelope modulation formats, a simple, sensible constraint to impose on the quadratic optimization is to confine the solution vector to lie within a sphere of radius \(|\mathbf{d}|\) passing through all possible data points. This is known in the area of optimization as a sphere-constrained quadratic program (SQP) [14]. The SQP problem has been studied intensively in the past and many results exist [15]. The solution for CDMA, following the satisfaction of the Karush–Kuhn–Tucker (KKT) conditions, is a linear detector which is closely related to the MMSE detector. This case has been considered independently in [16]–[18].

Constraining the data estimate vector to lie within a hypercube described by the data points leads to a problem which is known as a box-constrained quadratic program (BQP) [14]. Specifically, we consider the case where each element of the data estimate vector must lie in the range \([-1, 1]\). Again, this problem has been considered independently in [16] and [17], [19]. Such a problem is closely related to the linear complementarity problem (LCP) [20]. The LCP is equivalent to a BQP where each element of the data estimate vector is confined to \([0, 1]\). Both problems in turn are equivalent to special cases of the linear variational inequality (VI) problem over a rectangle [20]. A general iterative algorithm for solving the LCP was suggested by Ahn in [21]. This algorithm however, is based on the solution of the more general VI problem which includes the constrained quadratic optimization problem under consideration, i.e., the results in [21] can be extended to include the case of convex quadratic optimization subject to a CCS constraint.²

The algorithm can be interpreted as a general interference cancellation structure where some known successive and parallel approaches are recognized as special cases. The tentative decision function which plays an important role in interference cancellation is shown to be equivalent to an orthogonal projection onto the constraining space. It is then shown that the data estimate vector resulting from the clip-function interference cancellation schemes in [23] and [24] is asymptotically equivalent to the solution of the BQP, i.e., their performance approach that of the BQP detector as the number of stages tends to infinity. This interpretation also reveals that the projections for both the UQP and the SQP correspond to linear (soft) tentative decision functions, however, the latter case has a gradient which depends on the received signal. Moreover, we derive conditions for convergence for the general iterative algorithm which in turn provides new insight into convergence issues for some known nonlinear cancellation structures. Numerical examples show that only a few stages are needed for practical convergence.

At stated previously, at convergence the clip-function interference cancellation scheme provides an ML solution which is constrained to lie within a hypercube defined by the valid data points. The interference cancellation schemes based on a hyperbolic tangent function as a tentative decision suggested in [25]–[28] also provide a solution restricted to lie within the same hypercube. This solution is based on MMSE-optimal decision feedback [29] and is suggested in [27] as an iterative algorithm to find the nonlinear MMSE solution to the detection problem. It is shown that the true solution to this problem is a fixed point but convergence of the suggested algorithm is not proven. Simulation results indicate however that it does converge in most cases. This problem requires a thorough investigation which however, falls beyond the scope of the present paper.

The paper is organized as follows. In the following section, the uplink model is described. In Section III we discuss the solution of the constrained ML problem subject to a CCS constraint, introduce an iterative algorithm for solving it and consider convergence behavior. The relationship to known cancellation structures are presented in Section IV, and numerical examples are presented in Section V. Summarizing remarks concludes the paper in Section VI.

Throughout this paper scalars are lowercase, vectors are boldfaced lowercase, and matrices are boldfaced uppercase. The symbols \((\cdot)^T\) and \((\cdot)^{-1}\) are the transposition and inversion operators, respectively. All vectors are defined as column vectors with row vectors represented by transposition and \(\mathbb{R}\) denotes the set of real numbers. The notation \(\mathbb{D}^K\) denotes the set of all \(K\)-tuples over the set \(\mathbb{D}\), represented as column vectors.

### II. System Model

In this section, we describe the baseband uplink model of the CDMA communication system used throughout this paper. The uplink model is based on an asynchronous CDMA system with single-path channels and the presence of additive white Gaussian noise (AWGN) with zero mean and variance \(\sigma^2 = N_0/2\). In all our discussions, we assume that \(K\) users are active simultaneously.

User \(k\), \(k = 1, \ldots, K\), in this multiuser CDMA system transmits a binary information symbol stream \(d_k(n) \in \{-1, 1\}\) at data rate \(1/T_d\) symbols per second, where \(n = 0, 1, \ldots, L - 1\) is the symbol interval index and \(L\) is the length of the data block. To spread the signal, \(d_k(n)\) is modulated by a spreading waveform generated from a binary spreading code with a chip

²Algorithms for special cases of the VI problem have also been presented in [22].
rate of $1/T_c = N/T_d$ chips per second. The spreading code used to
modulate the $r$th bit can be written in a vector as $s_k(n) = (s_k(nN+1), s_k(nN+2), \ldots, s_k((n+1)N))^T$ with $s_k(j) \in \{(1/\sqrt{N}), (1/\sqrt{N})\}^N$. Binary data and chip formats are assumed for clarity only, and all the later concepts generalize to $M$-ary formats.

In addition, when the users are allowed random access to the channel, each user encounters a transmission delay relative to other users. The delay is measured against an arbitrary reference selected such that all the transmission delays are constrained to $0 \leq t_k < T_d$. Assume further that all the delays are constrained to be integer multiples of the chip interval. The relative delay normalized to the chip interval is then $\tau_k = t_k/T_c \in 0, 1, \ldots, N-1$. For a delay of $\tau_k$, the resulting transmitted discrete-time baseband signal due to symbol interval $n$ for user $k$ is then

$$a_k(n) = d_k(n) \left( \begin{array}{c} 0_{nN+\tau_k} \\ s_k(n) \\ 0_{(L-n)N-N-\tau_k-1} \end{array} \right) = d_k(n)c_k(n).$$

The dimension of $a_k(n)$ and $c_k(n)$ is $N_s = (L+1)N - 1$.

The continuous-time received signal is down-converted to baseband, passed through a filter matched to the chip pulse-shape and sampled. The received baseband signal after chip-matched filtering is then

$$r = \sum_{n=0}^{L-1} \sum_{k=1}^{K} d_k(n)c_k(n) + n$$

where $n$ is an $N_s$-dimensional vector of independent, identically distributed Gaussian random variables of zero mean and variance $\sigma^2 = (N_0/2)$. Here, we have assumed perfect power control and perfect phase synchronization for clarity and notational simplicity. The results do also apply to the general cases of no or arbitrary power control as well as to the case of randomly distributed received signal phases. A more convenient form of (1) is

$$y = S^T r = S^T S d + S^T n = Rd + \eta \in \mathbb{R}^{LK}$$

where $S$ is the matrix of received spreading codes and $d$ is the data symbol vector. This model is discussed in more detail in [30]. A minimal set of sufficient statistics of dimension $L K$ is obtained through correlation matched to the received spreading codes of the users. This is to ensure the maximization of the signal-to-noise ratio [31] and corresponds algebraically to

$$y = S^T r = S^T S d + S^T n = Rd + \eta \in \mathbb{R}^{LK}$$

where $y$ is the matched filter output vector, $R$ is the correlation matrix, and $\eta$ is a zero mean Gaussian noise vector with covariance matrix $\sigma^2 R$. In [12], it was shown that $R$ is symmetric positive definite (SPD) with probability one in an asynchronous system with arbitrary time delays, i.e., a chip asynchronous system. In the chip synchronous model defined here, there is a nonzero probability that $R$ is semi-positive definite. As $N$ increases, this probability diminishes. In the rest of the paper, we assume that $R$ is SPD.

III. CCS-CONSTRAINED ML DETECTION

The ML sequence detection criterion is defined as

$$u = \arg \max_{d \in \mathcal{D}} p(y|d)$$

which incidentally is identical to the MAP criterion assuming that all data symbols are equally likely. Here, $\mathcal{D}^{LK}$ represents the set of vectors in which the data estimate vector $u$ is assumed to exist. Since we are considering an AWGN channel, the negative log-likelihood function based on the $p(y|d)$ is described as

$$F(d) = \frac{1}{2} d^T R d - d^T y.$$ 

The general constrained ML problem for asynchronous CDMA is then described as

$$u = \arg \min_{\mathcal{V} \in \mathcal{D}^{LK}} F(v) = \arg \min_{\mathcal{V} \in \mathcal{D}^{LK}} \frac{1}{2} v^T R v - v^T y.$$ 

The so-called optimal ML detector for CDMA [2] is a $(0, 1)$-constrained minimization of $F(v)$ where $\mathcal{D} = \{-1, 1\}$, i.e., the ML solution is confined to $u \in \{+1, -1\}^{LK}$. The complexity in solving this problem is of the order of $2^{K-1}$, and thus grows exponentially with the number of users (see, e.g., [32] for details).

Here we relax the constraint to a CCS in order to limit the complexity of the solution algorithm, i.e., $\mathcal{D}^{LK} = C$. As special cases of $C$, we consider the real vector space $\mathbb{R}^{LK}$, a sphere determined by $\mathcal{S} = \{v \in \mathbb{R}^{LK} : |v|^2 \leq L K\}$, and finally a hypercube (box) described by $\mathcal{B} = \{v \in \mathbb{R}^{LK} : |b_j| < v_j \leq b_j \}$.

Here, $b = [1, \ldots, 1]^T$ is an $L K$-vector of all ones and $v \leq b$ is a compact notation for $v_j \leq b_j$ for all $j = 1, 2, \ldots, L K$, i.e., each element of $v$ is less than or equal to the corresponding element in $b$. We express this set in a compact way as $[-b, b]$. A CCS is element-wise separable (ES) if $\alpha_j \leq v_j \leq \beta_j \forall j = 1, 2, \ldots, L K$ where $\alpha_j$ and $\beta_j$ are appropriate constants. The corresponding constraint for element $j$ is denoted by $C_j$. Clearly $\mathbb{R}^{LK}$ and $\mathcal{B}$ are ES while $\mathcal{S}$ is not.

All the cases can be solved by satisfying the KKT conditions using Lagrange multipliers. The Lagrangian function associated with the SQP problem is

$$L(v, \lambda) = \frac{1}{2} v^T R v - v^T y + \frac{1}{2} \lambda (|v|^2 - L K).$$

A KKT point for the SQP is then described by a pair $(u, \lambda)$ [15] for which

$$(R + \lambda I) u = y$$

$$\lambda (|u|^2 - L K) = 0$$

$$\lambda \geq 0, \quad |u|^2 \leq L K.$$

It is well known that it is possible to completely detail the global solution of (3) under a sphere constraint without requiring any convexity assumption on the objective function.

**Proposition 1 (Proposition 2.1 [15]):** A point $u$ such that $|u|^2 \leq L K$ is a global solution of (3) if and only if there exists a unique $\lambda \geq 0$ such that the pair $(u, \lambda)$ satisfies the KKT condition in (4), and the matrix $(R + \lambda I)$ is positive semidefinite. If $(R + \lambda I)$ is positive definite, then (3) has a unique global solution.

For a proof, see [15]. In our case, we have assumed that $R$ is symmetric positive definite, and therefore there exists a
unique solution of the form $v = (R + \lambda I)^{-1}y$. This is identical in form to the MMSE solution which is described as $v = (R + \sigma^2 I)^{-1}y$ [33]. The two linear filters are not identical as the SQP detector confines the solution to be within a sphere while no such constraints are imposed on the MMSE solution. It can be shown however that $E[||v_t||^2] \leq LK$, so on average the MMSE detector does impose the same restriction [34]. Numerical examples have also revealed that $\lambda$ provides, on average, an accurate estimate of the noise variance $\sigma^2$. Furthermore, numerical examples show that the bit-error rate (BER) performance of the two detectors are virtually identical. With caution, it can therefore be claimed that there is a close relationship between the linear MMSE criterion and the sphere-constrained ML criterion. The unconstrained case follows from (4) by setting $\lambda = 0$. The result is the well-known decorrelator. No useful interpretation results from the satisfaction of the KKT conditions for the box-constraint.

A. Iterative CCSQP Detector

Since $F(v)$ is a strictly convex function as long as $R$ is SPD, the problem of minimizing $F(v)$ over a continuous region can be solved by a polynomial-time algorithm [35]. One approach to finding such an algorithm is to consider a VI problem defined as follows,

Definition 1: The VI Problem $VI(g, C)$ is defined as finding a vector $v \in C$ such that

$$(v - u)^T g(u) \geq 0 \quad \forall v \in C \quad (5)$$

where $g$ is a given continuous function from $C$ to $\mathbb{R}^{LK}$ and $C$ is a given CCS.

Let us further define

$$f(v) = \frac{\partial F(v)}{\partial v} = Rv - y.$$

We are then ready to establish a close link between CCS-constrained quadratic optimization and a special case of the VI problem.

Proposition 2: If $F(v)$ is a convex function and $u$ is a solution to $VI(f(v), C)$, then $u$ is a solution to the optimization problem in (3).

Proof: Since $F(v)$ is convex

$$F(v) \geq F(u) + (v - u)^T f(u) \quad \forall v \in C. \quad (6)$$

But $(v - u)^T f(u) \geq 0$, since $u$ is a solution to $VI(f(v), C)$. Therefore, from (6), it can be concluded that $F(v) \geq F(u)$ for all $v \in C$, i.e., $u$ is a minimum point of (3).

The above proposition is also true when $F$ is strictly convex. The proof follows the same arguments. So we can solve (3) by solving (5) with $g(v) = f(v) = Rv - y$.

The existence and uniqueness of a general solution to (5) follows directly from the proof of Proposition 2 and is summarized in the following proposition.

Proposition 3: There exists precisely one solution $u$ to (5) when $g(u) = f(u)$ if $R$ is positive definite.

Proof: If $R$ is positive definite, then $F(v)$ is a strictly convex function. From (3) it follows that $F(v) > F(u) + (v - u)^T f(u)$ for all $v \in C \setminus \{u\}$ and therefore $F(v) > F(u)$ for all $v \in C \setminus \{u\}$.

Having established the existence and uniqueness of a solution to (5), we move on to the critical question of how to find it. Before presenting the iterative algorithm, we need to define the orthogonal projection operation $P_C$ onto a CCS $C$.

Lemma 1: Let $C$ be a CCS in $\mathbb{R}^{LK}$. Then for each $y \in \mathbb{R}^{LK}$ there is a unique point $x \in C$ such that $||y - x|| \leq ||y - z||$ for all $z \in C$, and $x$ is known as the orthogonal projection of $y$ onto $C$. Using the fact that the projection is ESP, we can further state this corollary to Lemma 1.

Corollary 1: If $C$ is ES, then $x = P_C(y)$, for all $y \in \mathbb{R}^{LK}$, $x \in C$, and $x$ is a minimum point of $F(y)$.

Proof: Suppose that $u = P_C(u - \omega f(u))$, for some or all $\omega \geq 0$, and adding $(v - u)^T f(u)$ to both sides of the resulting inequality, we get

$$(v - u)^T f(u) \geq (v - u)^T f(u) - (u - \omega f(u)) \quad \forall v \in C. \quad (8)$$

From Lemma 2, it can be concluded that $u = P_C(u - \omega f(u))$, for some or all $\omega \geq 0$. Conversely, if $u = P_C(u - \omega f(u))$ for $\omega > 0$, then (8) directly and therefore (5) also holds true.

For a CCS constraint $C$ and $f(u) = R(u - y)$, the solution satisfies $u = P_C(u - \omega f(u))$ for some $\omega > 0$. An ES CCS, we can further state the following corollary to Lemma 3. Here we let $f_i(\cdot)$ denote element $i$ of the vector $f(\cdot)$.

Corollary 2: If $C$ is further ES then $u$ is a solution to $(v - u)^T f_i(u) \geq 0$, for all $v \in C$, if and only if $u = P_C(u - \omega f_i(u))$, for some or all $\omega \geq 0$, and any positive diagonal matrix $E$.

Proof: Since the constraining set is ES, the orthogonal projection is an ESP and then (7) can be restated as

$$u_i - u_i f_i(u) \geq 0 \quad \forall v_i \in C_i; \quad i = 1, 2, \ldots, LK.$$ (9)

The inequality in (9) can clearly be multiplied with any positive element $E_i$ and still be true. Again following Lemma 3, $(v_i - u_i) E_i f_i(u) \geq 0$ for all $v_i \in C_i$, $i = 1, 2, \ldots, LK$ if and only if $u_i = P_{C_i}(u_i - \omega E_i f_i(u))$ for some or all $\omega \geq 0$, $E_i \geq 0$, and $i = 1, 2, \ldots, LK$. Using the fact that the projection is ESP, we collect all the elements in the vector $u$ and the corollary is proven.
It therefore follows that for a ES CCS constraint \( C \), the solution satisfies
\[
\mathbf{u} = P_{C} [\mathbf{u} - \omega \mathbf{E} (\mathbf{R} \mathbf{u} - \mathbf{y})] \tag{10}
\]
for some \( \omega \geq 0 \) and any positive diagonal matrix \( \mathbf{E} \).

The following iterative algorithm is proposed in [21] for solving a linear complementarity problem. However, such a problem is a special case of the VI problem considered here [20] and the suggested algorithm is based on the above results for the general VI problem. The algorithm is thus also applicable in solving (10) for given \( \omega \) and \( \mathbf{E} \).

Algorithm 1: For any initial value \( \mathbf{u}_0 \in \mathbb{C} \), let
\[
\mathbf{u}_{m+1} = \mu P_{C} [\mathbf{u}_m - \omega \mathbf{E}_m (\mathbf{R} \mathbf{u}_m - \mathbf{y} + \mathbf{H} (\mathbf{u}_{m+1} - \mathbf{u}_m))] + (1 - \mu) \mathbf{u}_m \tag{11}
\]
where \( 0 < \mu \leq 1, \omega > 0 \) and \( m = 0, 1, \ldots, M - 1 \) is the iteration index. If the orthogonal projection operation can be decoupled into independent element-wise projections, then \( \mathbf{H} \) can be either strictly lower triangular, strictly upper triangular or equal to the null matrix \( \mathbf{O} \) and \( \mathbf{E}_m \) is any positive diagonal matrix. Otherwise, \( \mathbf{H} \) is equal to the null matrix \( \mathbf{O} \) and \( \mathbf{E}_m = \mathbf{I} \).

Algorithm 1 has the form of generalized interference cancellation. In fact, as shown later on for each of the three specific CCSs considered, special cases of the algorithm correspond to known successive and parallel cancellation structures. The algorithm is serial (successive) if nature when \( \mathbf{H} \) is strictly upper or lower triangular. Assuming that \( \mathbf{H} \) is lower triangular, the iteration of \( \mathbf{u} \) may be conducted element by element. When \( \mathbf{H} = \mathbf{O} \), the algorithm becomes parallel in nature, as iteration \( m \) only depends on estimates from iteration \( m - 1 \). This special case of the algorithm was first suggested in [36].

The choice of \( \mathbf{H} \), as well as \( \mathbf{E}_m \), \( \omega \) and \( \mu \) influence the convergence of the algorithm. Some useful sufficient conditions for convergence of the serial and parallel forms of (11) can however be found. Consider first the serial case. Let \( \mathbf{R} \) be partitioned such that \( \mathbf{R} = \mathbf{D} + \mathbf{L} + \mathbf{U} \), where \( \mathbf{D} \) is a diagonal matrix and \( \mathbf{L} \) and \( \mathbf{U} \) are strictly lower and upper triangular, respectively. Also, let \( \mathbf{E}_m = \mathbf{E} \).

**Theorem 1:** If \( \mathbf{H} \) is either \( \mathbf{L} \) or \( \mathbf{U} \), the sequence \( \{\mathbf{u}_m\} \) of Algorithm 1 with
\[
\mu \omega < 2 / \max_{k,n} (D_k(n)E_k(n)) \tag{12}
\]
where \( D_k(n) \) and \( E_k(n) \) are the diagonal elements of \( \mathbf{D} \) and \( \mathbf{E} \), respectively, corresponding to user \( k \), symbol interval \( n \), converges to the solution of the CCSQP for all realizations of \( \mathbf{d} \) and \( \mathbf{n} \).

**Proof:** The proof is based on showing that for all realizations of \( \mathbf{d} \) and \( \mathbf{n} \), \( \mathbf{u}_m \to \mathbf{u} \) for \( m \to \infty \) where \( \mathbf{u} \) denotes the unique fixed point. The complete proof is included in the Appendix.

When \( \mathbf{H} = \mathbf{O} \), the detector becomes a multistage parallel interference canceler, and the conditions for convergence change. Again, let \( \mathbf{E}_m = \mathbf{E} \) to allow for a convergence proof.

**Theorem 2:** If \( \mathbf{H} = \mathbf{O} \), the sequence \( \{\mathbf{u}_m\} \) of Algorithm 1 with
\[
\mu \omega < 2 / (A_{\max}F_{\max}) \tag{13}
\]
converges to the solution of the CCSQP for all realizations of \( \mathbf{d} \) and \( \mathbf{n} \), where \( A_{\max} \) and \( F_{\max} \) are the maximum eigenvalue of \( \mathbf{R} \) and the maximum diagonal element of \( \mathbf{E} \), respectively.

**Proof:** The proof is based on a similar strategy as the proof of Theorem 1 and is included in the Appendix.

The choice of other triangular forms for \( \mathbf{H} \) leads to an array of schemes based on block iterations [37], [38]. Previous results show for the linear case that most restricted and slowest convergence is found for \( \mathbf{H} = \mathbf{O} \) while least restricted and fastest convergence is found for \( \mathbf{H} = \mathbf{L} \) or \( \mathbf{H} = \mathbf{U} \). Choices in between bridge the gap between the two extreme. It is difficult, if not impossible, to analytically determine optimal values for the parameters \( \mathbf{E}_m \), \( \omega \), and \( \mu \). General trends on convergence as these parameters are varied are studied through computer simulations in Section V.

**IV. RELATIONSHIP TO INTERFERENCE CANCELLATION**

As mentioned earlier, special cases of Algorithm 1 are equivalent to known successive and parallel interference cancellation structures. An \( M \)-stage successive interference cancellation (SIC) scheme is described as
\[
\mathbf{u}_{m+1} = P_{C} [\mathbf{D}^{-1}(\mathbf{y} - \mathbf{L} \mathbf{u}_{m+1} - \mathbf{U} \mathbf{u}_m)] \tag{14}
\]
for \( m = 0, 1, 2, \ldots, M - 1 \) where \( \mathbf{u}_0 = \mathbf{y} \). The relationship between the SIC and the general iterative algorithm is clear from substituting \( \mathbf{H} = \mathbf{L} \), \( \mathbf{E}_m = \mathbf{D}^{-1} \) and \( \mu = \omega = 1 \) into (11). In cases where \( \mathbf{D} \neq \mathbf{I} \), i.e., with relaxed or no power control, selecting \( \mathbf{E}_m = \mathbf{D}^{-1} \) corresponds to a normalization of the received amplitudes, adjusting the signal levels to the relevant constraint. The fact that (14) indeed is describing practical interference cancellation for asynchronous CDMA was demonstrated for the linear case in [28] and [38].

Similarly, the \( M \)-stage weighted parallel interference canceler (PIC) [39] is described as
\[
\mathbf{u}_{m+1} = P_{C} [\beta_m \mathbf{y} + (1 - \beta_m \mathbf{R}) \mathbf{u}_m] \tag{15}
\]
for \( m = 0, 1, 2, \ldots, M - 1 \), which can be derived from (11) by taking \( \mathbf{H} = \mathbf{O} \). \( \mathbf{E}_m = \beta_m \mathbf{L} \), \( \omega = 1 \) and \( \mu = 1 \). Again \( \mathbf{u}_0 = \mathbf{y} \).

Due to the close relationship to interference cancellation, the general iterative algorithm in (11) can be efficiently implemented when \( \mathbf{H} = \mathbf{L} \) or \( \mathbf{H} = \mathbf{O} \). The concept of an interference cancellation unit (ICU) as a general building block can be applied, allowing for systematic construction. The corresponding ICU for user \( k \) at stage \( m \) is shown in Fig. 1. Here, \( \mathbf{e}_{m,k}(n) \) denotes the residual received vector for user \( k \) at stage \( m \),
symbol interval $n$ and $\Delta e_{m,k}(n)$ denotes the resulting update of the residual vector. The residual vector is obtained as

$$e_{m,k}(n) = \begin{cases} r - S_k(n)u_{m+1,k}(n) - S_F(n)u_{m,k}(n), & \text{for } H = L, \\ r - Su_m(n), & \text{for } H = 0 \end{cases}$$

with $S = [S_k(n)\ | S_F(n)]$ being a partition of $S$ where $S_k(n)$ contains the columns of $S$ associated with symbols already processed at stage $m+1$ while $S_F(n)$ contains the remaining columns of $S$. Similarly, $u_{m+1,k}(n)$ is a vector containing the stage $m+1$ estimates up until user $k$, symbol interval $n$, while $u_{m,k}(n)$ contains stage $m$ estimates of the remaining data symbols. A general successive cancellation structure is then constructed by inter-connecting ICU blocks as shown in Fig. 2.

The beauty of this representation is that other cancellation structures can be realized simply by changing the inter-connections of the ICUs. This is demonstrated by the PIC structure shown in Fig. 3.

The only difference between the two structures is when and where the residual vector update is carried out. In this representation, the inherent detection delay difference between successive and parallel structures is quite clear. Other combinations of SIC and PIC as suggested in [40] and [41] can also be implemented following this approach.

A. Tentative Decision Function

The orthogonal projection operation, essential for the algorithm, clearly corresponds to the tentative decision function in a cancellation structure. For unconstrained ML detection $C = R^{LK}$, the projection is defined by $P_{R^{LK}}(x) = x$ which obvi-
Fig. 4. Examples of the tentative decision functions corresponding to the orthogonal projections. The dashed curve represents the linear function for the UQP, the dashed and the dotted curves illustrate the functions for the SQP with \( \alpha = 1 \) and \( \alpha = 1/2 \), respectively, while the solid curve is the clipped soft decision represented by the BQP projection.

Fig. 5. 2-D example of projection onto a sphere region.

Fig. 6. 2-D example of projection onto a box region.

The orthogonal projection operation can be decoupled into independent element-wise projections, \( I_{\text{BQP}}(x_i) = x_i \). This is of course a linear (soft) decision function with gradient \( \alpha = 1 \) as illustrated in Fig. 4 by the dashed line. When the constraining space is a sphere, the element-wise (nonseparable) projection is described by

\[
(I_{\text{SQP}}(x)) = \begin{cases} \alpha x_i, & \text{if } ||x_i||^2 > LK \\ x_i, & \text{if } ||x_i||^2 \leq LK \\ \end{cases} \]

(17)

where \( \alpha = \sqrt{LK}/||x|| \) and \( 0 < \alpha \leq 1 \). This projection operation is illustrated in Fig. 5 for a two-dimensional (2-D) example. It is clear from the above that \( \alpha \) is decreasing for increasing \( ||x|| > \sqrt{LK} \). It is also clear that the overall projection does not decouple into independent element-wise projections. The sphere-constrained ML detector can therefore only be implemented as a PIC structure with \( \mathbf{H} = 0 \) and with \( \mathbf{E}_m = \mathbf{I} \). Also, since the projection operation depends on an entire data packet, an unacceptable detection delay is encountered. This delay together with other practical problems can be avoided by using an appropriate block approach [38] in order to perform the orthogonal projections. This has been done in the numerical examples. The corresponding tentative decision function is illustrated in Fig. 4 for \( \alpha = 1 \) (the dashed curve) and for \( \alpha = 1/2 \) (the dotted curve).

For the box constraint, the orthogonal projection operation can again be decoupled into independent element-wise orthogonal projections. The \( i \)th element of the orthogonal projection vector is in this case

\[
I_{\text{B}}(x_i) = \begin{cases} x_i, & \text{if } -1 < x_i < 1 \\ -1, & \text{if } x_i \leq -1 \\ +1, & \text{if } x_i \geq 1 \\ \end{cases} = \min\{\max\{-1, x_i\}, 1\}. \]

(18)

This projection is illustrated in the 2-D example shown in Fig. 6. The function \( I_{\text{B}}(x_i) \) is incidentally identical to the clipped soft-decision function suggested in [23] and [24] for interference cancellation. This tentative decision function is shown in Fig. 4 as the solid curve. It was shown in [23] that the clip-function SIC has better BER performance than the linear and hard-decision SIC, and it is therefore subject to much practical interest [42], [43]. The clip-function weighted PIC also performs better than the linear and hard-decision weighted PIC [24].

The general weighted PIC scheme corresponds to various partial cancellation schemes suggested in the literature for either the soft (linear) or the clipped soft tentative decision function [28], [37], [39], [44]–[46]. Here we have provided conditions for convergence for the case of \( \mathbf{E}_m = \mathbf{E} \), i.e., the weights stay fixed for all stages, but not necessarily the same for each user. For the linear case it is well-known that \( \mathbf{E} = \mathbf{D}^{-1} \) corresponding to the Jacobi iteration provides faster convergence than \( \mathbf{E} = \mathbf{I} \) which is usually denoted as the Richardson iteration [47]. The same is observed through simulations for the nonlinear cases. For the general case of \( \mathbf{E}_m \), it is difficult to advice any analytical conditions for convergence. For the linear case conditions have been presented in [45] and [46], however no results are currently available for the nonlinear cases [48]. In [28], [39], and [48], some observations based on simulations are presented.

**B. Convergence Issues**

An interesting observation regarding convergence is that Theorems 1 and 2 are independent of the constraining CCS. Hence, the conditions for convergence of the cancellation schemes are the same regardless of the tentative decision function. Note however that different schemes do not experience the same convergence rate and the same resulting BER performance. The results
in Theorems 1 and 2 are known for the linear case [37], however, an important corollary to Theorem 1 is that the clip-function SIC represented by (14), where $E_k(n) = (D_k(n))^{-1}$ and $\mu = \omega = 1$, always converges. Indeed, the good convergence behavior of the SIC, whether linear, hard-decision or clip-function, relative to the PIC has been known through simulations for some time. However, the result presented here represents the first proof for the guaranteed convergence of the clip-function SIC. Similar to the SIC case, an important corollary to Theorem 2 is that the clip-function PIC represented by (15), where $E_k(n) = (D_k(n))^{-1}$ and $\mu = 1$, always converges when $\omega < 2/E_{\max}$. Again, this result represents the first sufficient condition for convergence of the clip-function PIC.3

V. NUMERICAL RESULTS

In this section we investigate the BER performance and the rate of convergence of the SQP and BQP detectors based on numerical examples. Randomly selected long spreading codes are used and two different scenarios are considered, a lightly loaded case with $K = 10$ and $N = 32$, as well as a more highly loaded case where $K = 24$ and $N = 32$. The detectors are designed with $E = \mathbf{I}$ since we assume perfect power control, i.e., $D = \mathbf{I}$. The $\omega$ and $\mu$ of the PIC is selected based on simulation results such that the fastest possible convergence is achieved. Here we consider convergence to be reached when the BER for $m$ and $2m$ iterations are the same which is not necessarily the same as having $u_m = \bar{u}$. The initial data estimate $u_0$ is always chosen as $P_k(y)$.

First, we consider the SQP detector. In Fig. 7 the BER of the PIC is depicted for both cases. The BER of the linear MMSE detector is included for reference. For the lightly loaded case with $K = 10$, $\mu\omega = 0.7$, three iterations were required for convergence. As can be seen in the figure, the performance of the SQP and the linear MMSE detector are identical. The same behavior is observed for the more highly loaded case with $K = 24$. Here, $\mu\omega = 0.6$ and 7 iterations are required for convergence. Again the performance of the SQP and the linear MMSE detector coincide.

The influence of the choice of $\mu$ and $\omega$ has been investigated for both $K = 10$ and $K = 24$. The number of iterations required for convergence is obtained while systematically varying both $\mu$ and $\omega$. An important observation is that the rate of convergence for the cases investigated depends only on the product $\mu\omega$ and not on the individual values of $\mu$ and $\omega$. It therefore seems appropriate to select $\mu = 1$. The results are summarized in Table I. Here we can see that fastest convergence is achieved for $\mu\omega \approx 0.7$ when $K = 10$ while $\mu\omega \approx 0.6$ when $K = 24$. In these cases, $\mu\omega$ should be chosen as close as possible to the upper limit detailed by (13). It is also clear that the convergence rate of the PIC structure is quite sensitive to the $K/N$ ratio.

In Figs. 8 and 9 the BERs of the clipped SIC and the clipped PIC are depicted. The BER of the MMSE detector and the single-user bound are used as benchmarks. Fig. 8 shows the

---

3 As made clear above, the concept of using a clipped soft-decision function is quite important. The same function has also been mentioned in [39] and [44] for PIC.
BER of the lightly loaded case, where the clipped SIC needs three stages to achieve convergence whereas the clipped PIC ($\mu \omega = 0.9$) requires five stages to converge. For a three-stage clipped PIC, its BER performance is better than the MMSE detector and thus also better than the SQP-PIC. It can be seen that both interference cancellation schemes perform better than the MMSE detector after convergence but they are still quite far from the single user bound.

When the system becomes more highly loaded, the convergence rate decreases. This can be seen in Fig. 9, with the clipped SIC now needing six stages to converge and the clipped PIC ($\mu \omega = 0.7$) requiring 17 stages. Similarly, a six-stage clipped PIC has BER performance which is better than the SQP-PIC and MMSE detectors.

Again, the influence of the choice of $\mu$ and $\omega$ has been investigated for both $K = 10$, $K = 24$ and for both PIC and
TABLE II

<table>
<thead>
<tr>
<th>( \mu \omega )</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>( K = 10 )</td>
<td>12</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td>( K = 24 )</td>
<td>30</td>
<td>25</td>
<td>18</td>
<td>17</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

TABLE III

<table>
<thead>
<tr>
<th>( \mu \omega )</th>
<th>0.6</th>
<th>0.8</th>
<th>1.0</th>
<th>1.2</th>
<th>1.4</th>
<th>1.6</th>
<th>1.6</th>
<th>2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>( K = 10 )</td>
<td>8</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>8</td>
<td>-</td>
</tr>
<tr>
<td>( K = 24 )</td>
<td>12</td>
<td>9</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>9</td>
<td>-</td>
</tr>
</tbody>
</table>

SIC. Again for the cases investigated, only the product of \( \mu \omega \) is of importance. The results are summarized in Tables II and III. Here we can see that for the PIC, fastest convergence is achieved for \( \mu \omega \approx 0.9 \) when \( K = 10 \) while \( \mu \omega \approx 0.7 \) when \( K = 24 \). The same conclusions as for the SQP apply. In the SIC case, the best choice of \( \mu \) and \( \omega \) is \( \mu \omega \approx 1.0 \) for \( K = 10 \) and \( \mu \omega \approx 1.2 \) for \( K = 24 \). In these cases, the limit provided by (12) does not give any indication of the choice for fastest convergence. Another important observation is that the SIC structure is only marginally sensitive to the system load as only one additional iteration is required for \( K = 24 \) as compared to \( K = 10 \). Furthermore, the choice \( \mu \omega = 1.0 \), i.e., no weighting performed, does not lead to significant performance degradation.

Comparing the convergence rate characteristics for the PIC structures of both the SQP and the BQP cases, it is observed that some disagreement occur. According to Theorem 2, the convergence conditions should be the same for both cases. The reason for these discrepancies is the approximating window used to accommodate the orthogonal projection for the SQP. This level of approximation influences the convergence behavior.

VI. CONCLUDING REMARKS

In this paper, we have introduced the CCS constrained ML detector for CDMA as a complexity-limiting alternative to the (0, 1)-constrained ML detector usually denoted the optimal detector. Three special cases were used as illustrating examples, the unconstrained, the sphere-constrained and the box-constrained cases. A general iterative algorithm was suggested to solve the CCS-constrained ML problem and general conditions for convergence were derived. Known successive and parallel cancellation schemes were recognized as special cases of the algorithm. In fact, the algorithm is a general interference cancellation structure which can efficiently be implemented using an interference cancellation unit as a general building block in a systematic construction. The defining orthogonal projection onto the constraining space was shown to be equivalent to the tentative decision function in a cancellation structure. Specifically, the unconstrained ML detector was shown to lead to a linear tentative decision function while a sphere-constrained also lead to such a function. In this case however, the gradient depends on the length of the projected vector. The box-constraint was shown to correspond to a clipped soft-decision function for making tentative decisions, a function known to work well in interference cancellation.

The numerical examples show that the number of stages required to achieve the solution of the constrained ML problems depends strongly on the \( K/N \) ratio for a PIC structure while an SIC structure is only slightly influenced by the load. For both lightly and highly loaded systems, only a few stages of the clipped SIC are sufficient to achieve the solution of the BQP. A choice of \( \mu \omega = 1 \) (i.e., no additional weighting as compared to conventional SIC) is always close to the best possible choice for the SIC examples considered while for the PIC, the best choice of \( \mu \omega \) decreases with increasing \( K/N \).

APPENDIX

PROOFS OF THEOREMS 1 AND 2

Proof: [Theorem 1]: Referring to the definition of \( F(\mathbf{u}) \) in (3), if \( F(\mathbf{u}) \) is a decreasing function of \( m \), then \( \lim_{m \to \infty} \mathbf{u}_m = \bar{\mathbf{u}} \) for all realizations of \( \mathbf{d} \) and \( \mathbf{u} \) is the CCSQP solution. We prove this property of \( F(\mathbf{u}) \) given \( \mu \omega \leq 2/\max_{k \in \mathbb{C}} (D_k(n)E_k(n)) \) as follows:

\[
F(\mathbf{u}_{m+1}) - F(\mathbf{u}_m) = \left[ \omega (\mathbf{R}_m \mathbf{u}_m - \mathbf{y})^\top (\omega \mathbf{E})^{-1} (\mathbf{u}_{m+1} - \mathbf{u}_m) + \eta (\mathbf{u}_{m+1} - \mathbf{u}_m)^\top \mathbf{R} (\mathbf{u}_{m+1} - \mathbf{u}_m) - \mu (\mathbf{u}_{m+1} - (1 - \mu) \mathbf{u}_m) / \mu \right. \\
- \left. \left( (\mathbf{u}_m - \omega (\mathbf{R}_m \mathbf{u}_m - \mathbf{y}) + \mathbf{H} (\mathbf{u}_{m+1} - \mathbf{u}_m)) \right)^\top \right] (\omega E)^{-1} [ (\mathbf{u}_{m+1} - (1 - \mu) \mathbf{u}_m) / \mu - \mathbf{u}_m ] \\
+ (\mathbf{u}_{m+1} - \mathbf{u}_m)^\top \left( \frac{1}{2} \mathbf{R} - (\mu \omega (\mathbf{E})^{-1} - \mathbf{H}) \right) \times (\mathbf{u}_{m+1} - \mathbf{u}_m). \tag{19}
\]

We have that \( (\mathbf{u}_{m+1} - (1 - \mu) \mathbf{u}_m) / \mu \) is the CCSQP solution. It is then easy to show that \( \mathbf{u}_m \mathbf{u}_m - \mathbf{y} + \mathbf{H} (\mathbf{u}_{m+1} - \mathbf{u}_m) \). Hence, (19) may be rewritten as

\[
F(\mathbf{u}_{m+1}) - F(\mathbf{u}_m) = \mu [P_\mathbf{e} (\mathbf{e}_m) - \mathbf{e}_m]^\top (\omega \mathbf{E})^{-1} [P_\mathbf{e} (\mathbf{e}_m) - \mathbf{e}_m] + (\mathbf{u}_{m+1} - \mathbf{u}_m)^\top \left( \frac{1}{2} \mathbf{R} - (\mu \omega (\mathbf{E})^{-1} - \mathbf{H}) \right) \times (\mathbf{u}_{m+1} - \mathbf{u}_m). \tag{20}
\]

where \( \mathbf{e}_m = \mathbf{u}_m - \omega (\mathbf{R}_m \mathbf{u}_m - \mathbf{y}) + \mathbf{H} (\mathbf{u}_{m+1} - \mathbf{u}_m) \). But it is clear that each element of \( [P_\mathbf{e} (\mathbf{e}_m) - \mathbf{e}_m]^\top (\mathbf{E})^{-1} [P_\mathbf{e} (\mathbf{e}_m) - \mathbf{e}_m] \) is less than or equal to zero, because:

1) if \( \mathbf{e}_m \in \mathbb{C} \), \( P_\mathbf{e} (\mathbf{e}_m) - \mathbf{e}_m = 0 \); 2) if \( \mathbf{e}_m \not\in \mathbb{C} \), the corresponding elements of \( P_\mathbf{e} (\mathbf{e}_m) - \mathbf{e}_m \) and \( P_\mathbf{e} (\mathbf{e}_m) - \mathbf{e}_m \) have opposite signs because \( \mathbf{u}_m \in \mathbb{C} \). Therefore, we can conclude that

\[
(\mathbf{u}_{m+1} - f(\mathbf{u}_m) = \sum_{k=0}^{K-1} \sum_{n=0}^{L-1} \frac{1}{2} D_k(n) - (\mu \omega E_k(n))^{-1} \times (\mathbf{u}_{m+1,k}(n) - \mathbf{u}_m,k(n))^2 \leq 0 \tag{21}
\]

since \( \mathbf{R} \) is symmetric and \( \mu \omega \leq 2/\max_{k \in \mathbb{C}} (D_k(n)E_k(n)) \). Hence, the sequence \( \{F(\mathbf{u}_m)\} \) is nonincreasing. In fact, it is a
convergent sequence since $f$ is continuous on $\mathbb{C}$ and is thereby bounded. It follows then that $\lim_{m \to \infty} (F(u_{m+1}^T) - F(u_m)) = 0$. Now let $\overline{u}$ be an accumulation point, i.e., $\lim_{m \to \infty} u_m = \overline{u}$ of the sequence $\{u_m\}$. We show that $\overline{u}$ is in fact a solution of the CCSQP by noting that

$$0 = \lim_{m \to \infty} (F(u_{m+1}) - F(u_m)) \leq \lim_{m \to \infty} \sum_{j=1}^{K} \left( \frac{1}{2} D_j(k) - (\mu_j E_j(k) - 1)^2 \right) \times (u_{m+1,j} - u_{m,j})^2 \leq 0$$

and therefore $\lim_{m \to \infty} \Delta u_{m+1}^T = 0$ or $\lim_{m \to \infty} u_{m+1} = \lim_{m \to \infty} u_m = \overline{u}$. Hence, we have in the limit as $m \to \infty$, $PC[u_{m} - \omega E[R(u_{m} - y) + H(u_{m} - u_{m})] = u_{m}$ or $PC[u_{m} - \omega E[R(u_{m} - y)] = \overline{u}$, which according to Lemma 3 shows that $\overline{u}$ is the one and only solution of the CCSQP.

Proof: [Theorem 2]: In this case, the proof is identical (with $H = 0$) to the proof above until expression (21). We continue from there by defining $\Delta u_{m+1} = u_{m+1} - u_m$, and expressing $R$ in terms of its eigendecomposition as $VAV^T$, we then have

$$\left(u_{m+1} - F(u_m)\right) \leq \Delta u_{m+1}^T \left( \frac{1}{2} VAV^T - (\mu_j E_j) - \Delta u_{m+1} \right) \leq \frac{1}{2} A_{\text{max}} x_{m+1}^T - (\mu_j E_j) \Delta u_{m+1} \leq 0$$

where $x_{m+1} = V^T \Delta u_{m+1}$, since $\mu_j < 2/(A_{\text{max}} E_j).$

The remainder of the proof follows the proof above, with obvious substitutions.

REFERENCES


Peng Hui Tan (S’00) received the B. Eng. and M. Eng. degrees in electrical and electronic engineering from National University of Singapore in 1998 and 1999, respectively. He is currently working toward the Ph.D. degree within the Telecommunication Theory Group in the Department of Computer Engineering, Chalmers University of Technology, Gothenburg, Sweden.

Lars K. Rasmussen (M’93) was born on March 8, 1965, in Copenhagen, Denmark. He received the M.Eng. degree from the Technical University of Denmark, Lyngby, in 1989, and the Ph.D. degree from Georgia Institute of Technology, Atlanta, GA, in 1993.

From 1993 to 1995, he was at the University of South Australia, Adelaide, Australia. From 1995 to 1998, he was with the Centre for Wireless Communications at the National University of Singapore. He is currently an Associate Professor at Chalmers University of Technology, Gothenburg, Sweden.

Teng J. Lim (M’95) was born in Singapore on May 4, 1967. He received the B.Eng. degree in electrical engineering from the National University of Singapore in 1992 and the Ph.D. degree from Cambridge University, Cambridge, U.K., in 1995, in the area of IIR filtering for acoustic echo cancellation.

Since September 1995, he has been with the Centre for Wireless Communications, Singapore, where he is now a Senior Member of Technical Staff and leads the Signal Processing Strategic Research Group.