Abstract—Two computationally efficient algorithms are proposed for equalizing wideband MIMO channels in the Space-Time domain. Both of the proposed methods can be viewed as modifications to the Steepest Descent Method, for implementing the Zero-Forcing (or unconstraint Maximum-Likelihood) MIMO equalizer, which allow constraints about the feasible solution set to be incorporated in the iterative estimation process. The first approach assumes that the solution belongs in a closed convex set (inside a Hilbert space) and nearest point projection on the set is performed in each iterative step. The method is guaranteed to converge in the optimum point inside the convex set and hence provides significantly better performance than the standard linear equalizers. Fast convergence to a nearly optimal estimate can be achieved by the second proposed algorithm in which the new estimate in each iteration, is semi-projected on the discrete-discontinuous solution set. The proposed algorithms are of very low complexity as the number of arithmetic operations scales only linearly with the problem’s dimensionality.

I. INTRODUCTION

The importance of utilizing multiple antennas in order to provide high spectral efficiencies was illustrated in [1] and is already well appreciated by the telecommunications community. Over the last few years, a variety of Space-Time Coding (STC) architectures have been proposed, which can provide high spectral efficiencies as well as exploitation of the spatial diversity available due to the multiple receiving antennas. In this paper we focus on the simple uncoded spatial multiplexing architecture and propose new iterative methods for channel equalization in a wideband Single-Carrier (SC) MIMO system.

Optimal Maximum Likelihood (ML) sequence estimation in this system is practically impossible as the number of states in the Viterbi algorithm scales exponentially with the ISI channels’ memory and the number of transmit antennas. In SC wideband MIMO systems even block type linear equalizers (e.g. Zero Forcing (ZF) and MMSE) are not practical, as the number of complex multiplications required for the inversion of the channel matrix is related cubically to the information packet length and the square of the transmit antennas. This problem can be tackled by iterative methods for solving unconstraint quadratic optimization problems, such as the Steepest Descent Method (SDM) and the Conjugate Gradient Method (CGM) (see [2]). In these methods the computational complexity is dominated by matrix-vector multiplications. In the MIMO equalization problem the channel matrix has a block filtering matrix structure and therefore in these methods the number of complex multiplications (in each iteration) is only linearly related to the packet length.

This paper has two main contributions. Firstly an iterative method is proposed which converges to the Maximum-Likelihood (ML) estimate by constraining the solution to lie within a closed convex set, which is defined in a Hilbert space. The method, initially proposed in [3] and [4] for control problems, is known in the literature as the Gradient Projection Method (GPM). Here we make use of extended results which are provided by [5] and in particular the derivation of the GPM under mild differentiability conditions, which allow optimal step size parameter selection. Similar iterative algorithms for solving the Convex Constraint (CC) ML Multi-User Detection (MUD) problem in CDMA have been proposed in [9], [10]. These iterative algorithms ([9], [10]) coincide to previously proposed heuristic nonlinear MUD methods.

Furthermore a new iterative method is proposed which combines the SDM with the idea of semi-projecting the estimate on the discrete discontinuous solution set. More precisely, in each iteration of the SDM each soft estimate is checked to see whether it satisfies some adaptable threshold criterion. Symbols which satisfy the specified criterion are hard-decided while symbols which fail are left unchanged. Simulation results presented in section IV show that the proposed algorithm achieves the performance of the Probabilistic Data Association (PDA) algorithm (a state of the art suboptimal technique, which is proposed in [11] for MUD) but with two orders of magnitude less complexity within each iteration.

The remaining of this paper is organized as follows. In section II the physical layer of a SC spatially multiplexed MIMO system is modeled and the ST MIMO equalization problem is formulated. A brief reference is also made to standard linear equalization methods. In section III the SDM for implementing the ZF equalizer is described and the proposed algorithms...
are presented. Section IV provides simulation results for evaluating the BER performance of the two algorithms. Finally section V gives some conclusive remarks.

II. SYSTEM MODEL AND FORMULATION OF THE MIMO EQUALIZATION PROBLEM

A baseband, SC wideband MIMO system model with $N_t$ transmit and $N_r$ receive antennas is considered. An uncoded bit stream of length $L$ is subdivided into $N_t$ $(L/N_t)$-long substreams $\mathbf{a} = [a_1, \ldots, a_{N_t}]$. Each substream is modulated utilizing some modulation scheme of $k$ bits per symbol which is described by some discrete set of symbols $\Omega$ defined on the complex plane. The symbol streams $d_i, 1 \leq i \leq N_t$ are fed uncoded and simultaneously into all transmit antennas. The total transmit power $P$ is equally subdivided among the transmit antennas: $P = N_t$ utilizing some modulation scheme of $k$ duration of at least $T_d$.

The total transmit power $P$ is equally subdivided among the transmit antennas: $P = N_t$ per antenna. The duration of each modulated substream is $T_d = T_s(L/N_t)$ where $T_s$ is the symbol period. It is assumed that the fading characteristics of the propagation environment remain constant for a time duration of at least $T_d$. Each symbol substream follows a different propagation path to each of the receiving antennas and these paths are assumed to be highly uncorrelated. The frequency selective fading channels can be modeled in the time domain by FIR filters $h_{ij}, (1 \leq i \leq N_t, 1 \leq j \leq N_r)$ of memory order $M$. It is also assumed that guard-time intervals of duration $M_T$ are allowed between consecutive streams from each antenna. On the receiver end, each antenna ‘sees’ a linear combination of the channel-filtered transmitted subsequences contaminated also with spatially and temporally uncorrelated zero mean complex AWGN with variance of $\sigma^2/2$ per dimension. The received signal at receiving antenna $j$ is thus given by:

$$y_j^T = \sum_{i=1}^{N_t} (d_i * h_{ij})^T + v_j^T = \sum_{i=1}^{N_t} (H_{ij} d_i^T) + v_j^T = H_j d + v_j^T$$

(1)

where the operator (*) denotes discrete convolution between two sequences. $v_j \sim CN(0, \sigma^2 I)$ is a discrete realization of a complex Gaussian process and is of length $S = D + M = [L/(kN_t)] + M$. For brevity of representation and for aiding further analysis, the filtering effect of the sub-channel $h_{ij}$ is described by the channel filtering matrix $H_{ij}$ which is of dimensions $(S \times D)$ where each of its columns contains a downwards shift (relative to the previous column) of the vector $h_{ij}^T$. So in each column only $M + 1$ non zero elements exist. $H_j = [H_{ij1}, \ldots, H_{ijn}]$ and $d = [d_1, \ldots, d_{N_t}]^T$. We may concatenate the outputs of all receiving antennas in order to express a single linear system that describes compactly the input-output relationship of the MIMO system:

$$y = H d + v$$

(2)

where $y = [y_1, \ldots, y_{N_r}]$ of dimensions $(SN_r \times 1)$, $H = [H_1, \ldots, H_{N_r}]$ of dimensions $(SN_r \times DN_t)$ and $v = [v_1, \ldots, v_{N_r}]^T$. It is noted that $H$ has $(M+1)DN_rN_t$ non-zero entries. (2) provides the MIMO equalization problem which is the estimation of the transmitted information vector given the observation vector $y$ and typically also a good estimate of the system matrix $H$.

ML estimation of $d$ in (2) is practically impossible even in the case where the Viterbi algorithm is utilized as the number of states in the trellis increases exponentially with $(MN_t)$. Linear estimation methods such as Least Squares (LS) or MMSE estimation can be readily applied to the problem providing solutions:

$$\hat{d}_{LS} = (H^H H)^{-1} H^H y$$

(3)

$$\hat{d}_{MMSE} = (H^H H + \sigma^2 I)^{-1} H^H y$$

(4)

(3) is more widely known as the ZF equalizer as inter-signal and inter-antenna interference is forced to zero but at the expense of amplifying the noise in cases where $H$ is badly conditioned. (4) offers a fine balance between residual interference in the estimate and noise variance and generally provides a more reliable estimate than the ZF equalizer. Both approaches relax the constraint that the solution belongs in a discrete finite set $\Omega^{DN_r}$ and they solve the problem within $C^{DN_r}$. A direct implementation of the two equalizers would require the inversion of a matrix which is of dimensions $(DN_r \times DN_r)$ making the computational complexity cubically dependent to the packet length and thus impractical. A solution to this problem can be offered by methods which iteratively optimize convex functions, such as those optimized by the ZF and MMSE equalizers.

III. THE GRADIENT PROJECTION METHOD AND THE SEMI-PROJECTED STEEPEST DESCENT ALGORITHM FOR MIMO EQUALIZATION

It is easy to show that the ZF equalizer is the minimum to a quadratic optimization function:

$$W(d) = d^H R d - d^H z - z^H d + c$$

(5)

where $R = H^H H$ is an $N \times N$ complex symmetric matrix, $z = H^H y$ ($N \times 1$) is the MIMO matched-filter output and $c$ is some complex number. The direct solution to the problem is obtained by setting $\frac{\partial W}{\partial d} = 0$, $1 \leq n \leq N$ which leads to the following linear system to be solved: $R d_{opt} = z$. Obviously, direct inversion would provide the ZF equalizer. The SDM (see [2]) avoids the direct inversion of $R$ for obtaining $d_{opt}$ by assuming an arbitrary initial solution $d_o$ and moving towards the direction of steepest descent $-\nabla W$ which is parallel to $s_0 = z - R d_o$. So a new approximate solution is obtained as $d_1 = d_o + \gamma s_0$ , where $\gamma$ is a complex parameter whose optimal value can be determined locally by optimizing the quadratic $W(d_o + \gamma s_0)$ . In fact, $\gamma_{opt} = \frac{s_o^H s_o}{s_o^H R s_o}$ [2]. Because (5) is convex, the SDM guarantees convergence to the function’s unique minimum. The procedure is iterated until a suitable criterion is met. The SDM for implementing the ZF MIMO equalizer is given in Table I. It is emphasized that the block filtering matrix structure of $H$ allows the matrix-vector multiplications in the algorithm to be performed using linear filtering, making the number of
complex multiplications required within each iteration only linearly related to the packet size \( D \): \( O(IN^2 MD) \) where \( I \) is the number of iterations. This advantage also applies to the proposed methods.

A. Gradient Projection Method

In the case where the solution set is assumed to be closed convex, the GPM (see [3], [4], [5]) can be employed in order to iteratively optimize (5) by modifying the final step in the SDM as follows:

\[
\hat{d}_{i+1} = \Pi_Q(\hat{d}_i + \gamma_{opt} s_i)
\]

where \( \Pi_Q \) is the orthogonal projection operator on the convex set \( Q \). The convexity of the solution set as well as that of the optimization function, provide the existence of a unique solution to the CC-ML problem and this can be achieved by a polynomial time algorithm (see [9]). An important aspect of the GPM is the choice of the convex solution set. Existing work (see [9] and references within) has focused on constraining the solution to lie within a hyper-sphere or a hyper-cube. Adopting the hyper-cube constraint, the GPM can conveniently solve the CC-ML optimization problem: \( \min_d |y - Hd|^2 \) where \(-1 \leq d \leq 1, \forall n \) (assuming BPSK modulation) as follows (\( s_i \) and \( \gamma_{opt} \) are calculated as in Table I):

\[
\hat{d}_{n,i+1} = \begin{cases} 
-1, & \text{if } \text{real}(\hat{d}_{n,i} + \gamma s_{n,i}) < -1 \\
\hat{d}_{n,i} + \gamma s_{n,i}, & \text{if } -1 \leq \text{real}(\hat{d}_{n,i} + \gamma s_{n,i}) \leq +1 \\
+1, & \text{if } \text{real}(\hat{d}_{n,i} + \gamma s_{n,i}) > +1 
\end{cases}
\]

An attractive feature of the GPM, which is used in (7) is that the projection operation can be applied independently for each component of the solution, making the complexity indifferent compared to that of the SDM. Figure 1 illustrates the convex set projection idea for the examples of BPSK and QPSK modulation. The exact choice of the shape and size of the convex solution set will depend on the signaling constellation used. Care needs to be taken in choosing the convex region when complex signaling is employed since in many modulation schemes, information in the real and imaginary components cannot be treated independently. In such cases the convex projections cannot be applied separately for the real and imaginary components of the signal. Furthermore the GPM also requires the initial solution to lie within the assumed closed convex set.

### Table I

| \( \hat{d}_a \) (e.g. = 0) for \( i=1/I \) | \( s_i = (z - R^H \hat{d}_i) \) | \( \gamma_{opt} = \frac{	ext{TR}^H_z}{2} \) | \( \hat{d}_{i+1} = \hat{d}_i + \gamma_{opt} s_i \) |

---

B. Semi-Projected Steepest Descent Algorithm

A modification to the SDM, which can achieve better performance than the unconstraint ML and the CC-ML solutions as well as faster convergence without increase in the complexity, is further proposed. The structure of the proposed algorithm is similar to that of the GPM but it differs in the way the feasible solution set is defined. In particular, in each iteration estimated symbols which have converged sufficiently away from all signaling values but one are hard decided to the nearest signaling point. On the other hand symbols which fail the specified threshold criterion are left unaltered in their soft form.

In the algorithm each new estimate is a linear combination of random variables and thus the Central Limit Theorem (CLT) is assumed to hold. The assumption that each estimate closely follows a Gaussian (and therefore unimodal) distribution means that only the most reliable estimates are hard-decided in each iterative step of the algorithm. Significant amount of work has been carried out in MUD for CDMA systems in order to prove asymptotic Gaussianity of linear estimation methods [7], [8]. As the MIMO equalization problem is fundamentally the same we assume that asymptotic Gaussianity of the estimates is a valid assumption in it too. The Gaussianity assumption also allows to calculate posterior Log-Likelihood Ratios (LLR) for each estimated symbol (\( \tilde{d} \)) in each iteration:

\[
\Lambda_j = \frac{\ln(\omega_j | \tilde{d})}{\ln(\omega_0 | \tilde{d})} = \frac{\| \tilde{d} - \omega_j \|^2}{\| \tilde{d} - \omega_0 \|^2}
\]

where \( \omega_j \in \Omega \) (i.e. a point in the signaling constellation) and \( \omega_0 \) is the reference signaling point which lies closest (in the Euclidian distance sense) to \( \tilde{d} \). These posterior LLR can be used as a formal criterion for deciding whether estimates should be hard-decided or left unaltered. For modulation schemes where information on the real channel is independent to information on imaginary channel, the semi-projection can be performed independently for the real and imaginary components of the estimates (Figure 2). The specified threshold boundaries are chosen to be adaptable and in particular to converge towards the classical decision thresholds boundaries (in which case all symbols are hard decided). In this way
TABLE II

Semi-Projected Steepest Descent Algorithm with QPSK Signaling

\[
\tilde{a}_0 (\text{e.g. } = 0) \\
\text{for } i=0:1: \lfloor \frac{N}{2} \rfloor \\
\text{for } n=1:N (= DN_t) \\
\begin{cases} \\
\text{if } |\text{real}(\tilde{a}_{n,i})| > \Delta_i \\
\tilde{a}_{n,i} = \text{sign}(\text{real}(\tilde{a}_{n,i})) + j \cdot \text{imag}(\tilde{a}_{n,i}) \\
\text{else} \\
\tilde{a}_{n,i} = \text{real}(\tilde{a}_{n,i}) + j \cdot \text{imag}(\tilde{a}_{n,i}) \\
\text{if } |\text{imag}(\tilde{a}_{n,i})| > \Delta_i \\
\tilde{a}_{n,i} = \text{real}(\tilde{a}_{n,i}) + j \cdot \text{sign}(\text{imag}(\tilde{a}_{n,i})) \\
\text{else} \\
\tilde{a}_{n,i} = \text{real}(\tilde{a}_{n,i}) + j \cdot \text{imag}(\tilde{a}_{n,i}) \\
\end{cases} \\
\gamma_{\text{opt}} = \frac{\Delta_i}{2 \cdot \text{A}} \\
\tilde{a}_{i+1} = \tilde{a}_i + \gamma_{\text{opt}} s_i \\
\end{cases}
\]

the new estimate in each iteration is partially projected on the discrete solution set and thus the proposed algorithm is termed as Semi-Projected Steepest Descent Algorithm (SPSDA).

Table II provides the SPSDA when QPSK signaling is employed. \( \Delta_i \) in Table II represents the value of the adaptable threshold boundaries which are symmetrical around zero on the real and imaginary axes. In this example, the GPM can be viewed as a special case of the SPSDA if the hyper-cube convex constraint is adopted and the threshold boundaries are fixed to \( \Delta_i = 1, \forall i \).

Fig. 2. Semi-Projection idea in the SPSDA with QPSK signaling

IV. SIMULATION RESULTS

The proposed methods have been simulated for a spatially multiplexed uncoded SC wideband MIMO system. The following assumption have been made: \( N_t = N_r = 4, L = 1024, \) QPSK modulation where pairs of bits are mapped independently on the real and imaginary channels. Uncorrelated slow Rayleigh fading sub-channels (\( h_{ij} \)) with \( M = 3 \) and uniform delay profile have also been assumed. At the receiver perfect knowledge of the MIMO channel state has been assumed. Sufficiently long silent periods between sub-stream transmissions have been allowed so that interference from past and future sub-streams is avoided.

In the GPM each estimated symbol is assumed to lie within a rectangular closed region on the complex plane whose vertices are the 4 signaling points on the plane (as in the example given in Figure 2 but with \( \Delta \) fixed). Similar approach has been followed for the SPSDA but \( \Delta_i \) on the real and imaginary axes has been initialized to \( \Delta_i = 1.2(\sqrt{2}/2)A \) and programmed to decrease to \( \Delta_i = 0.2(\sqrt{2}/2)A \) in equal steps with each iteration, where \( A \) is the square root of symbol energy and \( I = 30 \). In order to enhance further the performance, this procedure was repeated for a few times (so the decision boundaries were reset and allowed to shrink again in the same manner) but the initial solution in each run was the estimate from the previous run. The simulation results have shown that 2 runs where sufficient to provide near optimal performance. In both methods an initial solution of 0 was assumed for all estimates. From Figure 3 it is observed that both of the proposed methods can provide both fast convergence and substantially superior performance than the standard linear equalization methods. The most attractive performance is provided by the SPSDA which achieves more than 10dB gain relative to the unconstrained solutions.

In a second simulation the performance of the SPSDA has been compared to this of the PDA algorithm, proposed in [11] for performing MUD in CDMA systems. The algorithm was shown to provide near optimal performance in a complexity cubically related to the problems dimensionality and is thus considered as a state of the art MUD algorithm. In this example BPSK modulation was assumed, \( L = 25 \) and \( N_t = N_r = 3 \). Real Rayleigh fading channels with \( M = 3 \) and uniform delay profile have also been assumed. The results given in Figure 4 show that in the simulated example, the SPSDA is capable of achieving the performance of the PDA algorithm. However in the SPSDA this is achieved in substantially less complexity. In particular the complexity of PDA is: \( O(I[MDN_t^2]) \) while the complexity of the SPSDA is: \( O(IMDN_t^2) \), i.e. two orders of magnitude less. Nevertheless the PDA algorithm requires only 2 or 3 iterations to achieve near optimal performance.

V. CONCLUSIONS

Two new iterative algorithms have been proposed for equalization in wide-band MIMO systems. The GPM provides a computationally efficient iterative solution to the well-defined convex constraint ML optimization problem. On the other hand the SPSDA is proposed as a new heuristic algorithm for MIMO equalization. Both algorithms provide significantly superior performance compared to classical suboptimal linear equalizers and this is achieved without an increase in computational complexity. The SPSDA in particular can achieve
the performance of the state-of-the-art PDA algorithm but in two orders of magnitude less complexity. It is also stressed that despite the seemingly high number of iterations involved in the proposed algorithms, their overall computational complexity is low because of the very small number of computations involved within each iteration.

ACKNOWLEDGEMENTS

The first author would like to thank QinetiQ Ltd. and EPSRC for funding his research at the University of Bristol.

REFERENCES