Multilayer spatial multiplexing in next-generation WiMAX

IIMC International Information Management Corporation, 2009
ISBN: 978-1-905824-12-0

Multilayer Spatial Multiplexing in next-generation WiMAX

Alberto TARABLE[1], Guido MONTORSI[2]
[1]*Politecnico di Torino, C.so Duca degli Abruzzi 24, Torino, 10129, Italy*
Tel: +39 011 0904228, Fax: +39 011 0904099, Email:alberto.tarable@polito.it
[2]*Politecnico di Torino, C.so Duca degli Abruzzi 24, Torino, 10129, Italy*
Tel: +39 011 0904144, Fax: +39 011 0904099, Email:guido.montorsi@polito.it

Abstract: We propose in this paper as a possible candidate for next-generation WiMAX a multilayer spatial multiplexing scheme, which is flexible and conceptually simple. Every layer is a QPSK symbol stream, weighted by a coefficient, and superimposed on the other layers. A method to optimize the coefficients is described in details. At the receiver, an iterative (turbo-like) structure gives very good performance at an affordable complexity, as it is shown by simulation results, in a realistic and WiMAX-complying scenario.

Keywords: WiMAX, multilayer modulation, spatial multiplexing

1. Introduction

Currently, the IEEE 802.16m standardization group is working toward an update of the IEEE 802.16e standards, in order to meet the ever-increasing requirements for speed, bandwidth, flexibility and offered services and functionalities. This new standard will support the second generation of broadband wireless access systems, after the popular one which is commonly known as WiMAX. One of the key challenges the requirements of this new generation poses is a more efficient use of the degrees of freedom available in the MIMO channel, while MIMO signalling techniques are currently limited to the alternative between the Alamouti scheme of [1] and a plain spatial multiplexing (SM) with two or four antennas [2].

In this paper, we introduce a SM scheme which is *multilayered*, together with an iterative (turbo-like) receiver, and we show that it has a good performance over a multipath channel which is well-suited to the WiMAX scenario.

Multilayer schemes have a rich history in the literature. Here, because of space limitations, we only cite [3], which shares with our scheme also the iterative receiver. However, in [3], a scheme is considered in which every symbol is repeated and transmitted over the different transmit antennas, in order to obtain the maximum possible diversity gain, while no multiplexing gain is achieved. Instead, we do not replicate symbols across antennas, but *demultiplex* symbol streams over the antennas, to obtain maximum multiplexing gain. As for the diversity gain, in our scheme it is delegated to a powerful channel code (the convolutional turbo code of the IEEE 802.16e standard) to exploit the intrinsic diversity of the multipath channel.

The structure of the paper is as follows. In Sect. 2., we describe the transmitter and the receiver of the proposed scheme. In Sect. 3., we describe a technique to optimize the powers of the different layers. In Sect. 4., we show some simulation results. In Sect. 5., finally, we draw some conclusions.

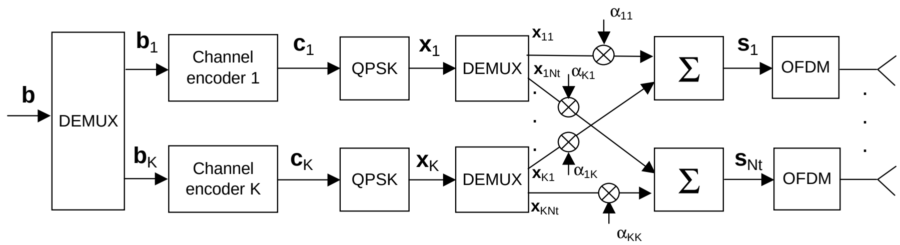

Figure 1: Block diagram of the transmitter for the proposed scheme.

2. Description of the multilayer SM scheme

2.1 The transmitter

Let K be the number of layers in the multilayer scheme. The step-by-step description of the operations performed at the transmitter side is now given. A block scheme of the transmitter is shown in Fig. 1.

- The binary input information word $\mathbf{b}$ is first segmented in a demultiplexer into K subwords $\mathbf{b}_1, \ldots, \mathbf{b}_K$, each of the same length K_c.
- In the k-th layer, the subword $\mathbf{b}_k$ enters a channel encoder, whose output binary codeword, denoted by $\mathbf{c}_k$, has length N_c. As it is evident from notation, we assume that all K channel encoders have the same input and output lengths, and thus also the same code rate $R_c = K_c/N_c$. In the following, it will be supposed that these K channel encoders are obtained from a fixed encoder, equal for all layers, cascaded with layer-specific interleavers.
- Codeword $\mathbf{c}_k$ is then modulated, according to a given modulation format, which will be QPSK throughout this paper. Let $\mathbf{x}_k$ be the stream of QPSK symbols obtained by modulating $\mathbf{c}_k$. We suppose without loss of generality that these QPSK symbols are normalized to have unitary energy, i.e., every symbol belongs to the set $\{(\pm 1 \pm j)/\sqrt{2}\}$.
- The stream $\mathbf{x}_k$ is input to a demultiplexer, which segments it into N_t substreams $\mathbf{x}_{k1}, \ldots, \mathbf{x}_{kN_t}$, all of the same length, where N_t is the number of transmit antennas. Substream $\mathbf{x}_{kj}$ is the contribution from layer k to transmit antenna j.
- Supersymbols for the j-th transmit antenna are then formed by superimposing QPSK symbols from the K layers in the following way:

$$\mathbf{s}_j = \sum_{k=1}^{K} \alpha_{kj} \mathbf{x}_{kj}, \tag{1}$$

  where the α_{kj} will be called *combination coefficients* afterwards, and are the key parameter to determine the performance of the overall system.
- At the j-th transmit antenna, $\mathbf{s}_j$, endowed with a cyclic prefix, is input to an OFDM modulator, in order to cope with the frequency selectivity of the channel. All N_t OFDM modulators are equal and use N_c subcarriers. The OFDM modulator output is then transmitted through the corresponding antenna.

The resulting spectral efficiency of the proposed scheme is given by $\eta = 2KN_tR_c$.

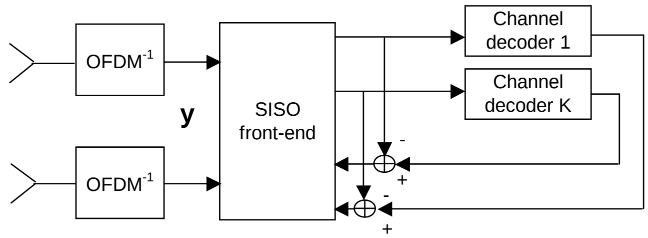

Figure 2: Block diagram of the receiver for the proposed scheme.

2.2 The receiver

The block scheme of the receiver for the K-layer scheme described in the previous subsection is shown in Fig. 2. The receiver is equipped with N_r receive antennas, with $N_r \geq N_t$ in our scheme, since we want to achieve a multiplexing gain equal to N_t. Let $\mathbf{y}[l]$ be the signal received at time step l from the receive antennas, after OFDM demodulation and removal of the cyclic prefix. If the time variations of the channel are slow enough w.r.t. the OFDM symbol duration, then $\mathbf{y}[l]$ can be written as:

$$\mathbf{y}[l] = \mathbf{H}[l]\mathbf{A}\mathbf{x}[l] + \mathbf{n}[l], \tag{2}$$

where

- $\mathbf{H}[l]$ is the $N_r \times N_t$ matrix of the realization at time step l of the equivalent frequency-domain flat-fading channel,
-
$$\mathbf{A} = \begin{bmatrix} \alpha_{11} & \dots & \alpha_{1K} & 0 & \dots & \dots & \dots & \dots & 0 \\ & & & & \ddots & & & & \\ 0 & \dots & \dots & \dots & \dots & 0 & \alpha_{N_t,1} & \dots & \alpha_{N_t,K} \end{bmatrix},$$
- $\mathbf{x}[l] = [x_{11}[l], \dots, x_{1K}[l], \dots, x_{N_t,1}[l], \dots, x_{N_t,K}[l]]^T$ is the vector of KN_t QPSK symbols from all layers used to form the transmitted supersymbols at time step l, and
- $\mathbf{n}[l]$ is a vector of uncorrelated Gaussian noise samples with zero mean and variance σ^2.

As it can be seen, the expression of $\mathbf{y}[l]$ can be interpreted as the output of a multiple access channel, over which KN_t users transmit simultaneously. Following this analogy, in order to decode the information, we use an iterative structure that exploits the turbo principle, akin to those proposed in [3]-[5]. This structure is composed by two main blocks, which exchange information along a certain number of iterations.

The first (inner) block has appeared in several incarnations in the literature, variously called Wang-Poor [4], elementary signal estimator [3], etc. In this work, we will simply call it single-input single-output (SISO) front-end. It is useful to suppress the inter-antenna interference as well as the inter-layer interference. Notice that the latter has been artificially introduced at the transmitter. The second (outer) block is a bank of K SISO channel decoders, one for each layer.

Let us consider the n-th iteration in the decoding process. The SISO front end has two inputs, $\mathbf{y}[l]$ and the feedback from the bank of SISO decoders, which was computed in the $(n-1)$-th iteration. We refer to the latter as the current *bit statistics*, in the sense that it is interpreted as an *a priori* knowledge on the codeword bits. From this bit statistics, supposed to be independent among different bits, a current *symbol* statistics can be derived, which is used by the SISO front-end as apriori information on the symbols emitted by each antenna and each layer.

At the n-th iteration, the SISO front-end performs three steps, which are described hereafter.

- *Soft interference cancellation*: the estimate, based on the current symbol statistics, of the inter-antenna and inter-layer interference on the symbol transmitted through antenna j by layer k is subtracted from $\mathbf{y}[l]$:

$$\mathbf{y}_{kj}^{(n)}[l] = \mathbf{y}[l] - \mathbf{H}[l]\mathbf{A}\widetilde{\mathbf{x}}_{kj}^{(n)}[l], \tag{3}$$

where

$$\left(\widetilde{\mathbf{x}}_{kj}^{(n)}[l]\right)_{k'j'} = \begin{cases} \widetilde{x}_{kj}^{(n)}[l], & (k',j') \neq (k,j), \\ 0 & \text{otherwise}, \end{cases} \tag{4}$$

and

$$\widetilde{x}_{kj}^{(n)}[l] \triangleq E\left[x_{kj}[l] | n\text{-th iteration}\right] \tag{5}$$

is the average of $x_{kj}[l]$, based on the current symbol statistics.

- *Linear MMSE filtering:* to reduce residual interference, $\mathbf{y}_{kj}^{(n)}[l]$ is linearly filtered in the following way:

$$\widetilde{y}_{kj}^{(n)}[l] = \mathbf{m}_{kj}^{(n)}[l]\mathbf{y}_{kj}^{(n)}[l], \tag{6}$$

where $\mathbf{m}_{kj}^{(n)}[l]$ is a row vector, solution of the following standard linear MMSE problem:

$$\mathbf{m}_{kj}^{(n)}[l] = \arg\min_{\mathbf{m}} E\left[|x_{kj}[l] - \mathbf{m}\,\mathbf{y}_{kj}^{(n)}[l]|^2\right] \tag{7}$$

and is explicitly given by ([6]):

$$\mathbf{m}_{kj}^{(n)}[l] = \mathbf{e}_{kj}^T(\mathbf{H}[l]\mathbf{A})^H\left(\mathbf{H}[l]\mathbf{A}\mathbf{V}_{kj}^{(n)}[l]\mathbf{A}^H\mathbf{H}[l]^H + \sigma^2\mathbf{I}\right)^{-1}, \tag{8}$$

where

$$\mathbf{V}_{kj}^{(n)}[l] = \mathbf{V}^{(n)}[l] - \mathrm{var}\left(x_{kj}[l]\right)\mathbf{e}_{kj}\mathbf{e}_{kj}^T \tag{9}$$

and

$$\begin{aligned}\mathbf{V}^{(n)}[l] = \mathrm{diag}\,(&\mathrm{var}\left(x_{11}[l]\right), \,...,\, \mathrm{var}\left(x_{1K}[l]\right), \,..., \\ &...,\, \mathrm{var}\left(x_{K1}[l]\right), \,...,\, \mathrm{var}\left(x_{KN_t}[l]\right)),\end{aligned} \tag{10}$$

is the diagonal matrix of variances for all symbols, based on the current symbol statistics.

- *Extraction of the output*: it consists in delivering to the bank of SISO decoders the a-posteriori symbol statistics from $\widetilde{y}_{kj}^{(n)}[l]$. Since the exact computation would be exponentially complex with the product KN_t, the Gaussian approximation (GA) is invoked, which consists in approximating the sum (residual interference + noise) in $\widetilde{y}_{kj}^{(n)}[l]$ with a Gaussian random variable with the same mean and variance:

$$\widetilde{y}_{kj}^{(n)}[l] \simeq \beta_{kj}^{(n)}[l] x_{kj}[l] + \nu_{kj}^{(n)}[l], \tag{11}$$

where $\nu_{kj}^{(n)}[l] \sim \mathcal{N}\left(\mu_{kj}^{(n)}[l], \rho_{kj}^{(n)2}[l]\right)$. The mean and the variance of $\nu_{kj}^{(n)}[l]$, as well as the value of $\beta_{kj}^{(n)}[l]$, can be easily computed. For the details, see [6].

The output of the SISO front-end is used to compute a-posteriori bit statistics. Those for the k-th layer are then grouped and sent to the k-th SISO decoder. In our scheme, the latter includes a layer-specific deinterleaver and a standard turbo decoder, which, after a predefined number of turbo iterations, provides soft information on the encoded bits, which will be used, after interleaving, to form the symbol statistics for the next iteration. In the last iteration, the SISO decoders must also supply hard estimates of the information bits, which represent the output of the whole receiver.

3. Optimization of the combination coefficients

In this section, we describe the method we used to set the combination coefficients for the schemes whose simulation results are shown in the next section.

We have imposed some constraints on the possible values of the α_{kj}'s, namely:

- $\alpha_{kj} = \alpha_k$, i.e., the combination coefficients do not depend on the considered transmit antenna, but are applied to a given layer as a whole, and
- α_k is real (and positive), which means scaling of the layers, but *not* rotations.

These constraints configure a multilayer scheme in which every layer is just normalized to a given power and superimposed on the other layers. This restriction, although not necessarily optimal in general, arises from the parallel between such scheme and a multiple access channel, in which different users are received at different power levels, due to the near-far effect.

The signal-to-noise ratio at the output of the SISO front-end for the k-th layer, averaged w.r.t. transmit antenna j and time step l, can be computed by using (6) and (11), together with (9) and the matrix inversion lemma ([3]):

$$\mathrm{SNR}_k^{(n)} = \frac{1}{N_t} \sum_{j=1}^{N_t} E_l \frac{\gamma_j^{(n)}[l] P_k}{1 - \gamma_j^{(n)}[l] v_{kj}^{(n)}[l] P_k}, \tag{12}$$

where $\gamma_j^{(n)}[l] = \left(\sigma^2 \left(\mathbf{H}[l]^H \mathbf{H}[l]\right)^{-1} + \mathbf{A}\mathbf{V}^{(n)}[l]\mathbf{A}^H\right)^{-1}_{jj}$, $v_{kj}^{(n)}[l] = \mathrm{var}\left(x_{kj}[l]\right)$ and $P_k = \alpha_k^2$.

We suppose that the channel decoder used is a threshold device: if the SNR at its input is below a given value SNR_{th}, the soft information delivered at its output is precisely zero, while if it exceeds SNR_{th}, then the output information is perfect. In this way, if at iteration n, $\mathrm{SNR}_k^{(n)} < \mathrm{SNR}_{th}$, at the $(n+1)$-th iteration, $v_{kj}^{(n+1)}[l] = v_{kj}^{(1)}[l] = 1$ (the energy of transmitted symbols) for all l and j, while if $\mathrm{SNR}_k^{(n)} > \mathrm{SNR}_{th}$, $v_{kj}^{(n+1)}[l] = 0$ for all l and j. Such an assumption is well fitted for powerful channel codes, like turbo codes or LDPC codes with large block lengths, while it is a coarse approximation for other kinds of codes, like convolutional codes, for instance.

Using this assumption, we can easily find the evolution of the SNRs for the different layers along the iterations. Starting from $v_{kj}^{(1)}[l] = 1$ for all k, j and l, (12) gives $\mathrm{SNR}_k^{(1)}$ for all k, from this we can find $v_{kj}^{(2)}[l]$ applying the above threshold condition, and so forth.

Let us order the layers so that $P_1 \geq P_2 \geq \cdots \geq P_K$. Consider iteration n, with $n \leq K$, and suppose that layers from 1 to $n-1$ have been decoded, while layers from n up to K have not yet been decoded. The SNR at the output of the SISO front-end for layer n (the strongest layer left) will be:

$$\mathrm{SNR}_n^{(n)} = \frac{1}{N_t} \sum_{j=1}^{N_t} E_l \frac{\gamma_j^{(n)}[l] P_n}{1 - \gamma_j^{(n)}[l] P_n} \geq \frac{\overline{\gamma}^{(n)} P_n}{1 - \overline{\gamma}^{(n)} P_n}, \tag{13}$$

where $\overline{\gamma}^{(n)} = \frac{1}{N_t} \sum_{j=1}^{N_t} E_l \gamma_j^{(n)}[l]$. The inequality derives from Jensen's inequality, applied to the convex function $f(x) = x/(1-x)$. The n-th layer will certainly be decoded if

$$\frac{\overline{\gamma}^{(n)} P_n}{1 - \overline{\gamma}^{(n)} P_n} \geq \mathrm{SNR}_{th}, \tag{14}$$

which is not solvable in closed form w.r.t. P_n, because $\overline{\gamma}^{(n)}$ depends on it. However, this is the basis for our optimization technique. Precisely, we start with the K-th layer (the weakest) and we solve numerically (14) with the equal sign, obtaining P_K, then we go to layer $K-1$ and we solve again (14), obtaining P_{K-1}, and so forth.

Notice that $\overline{\gamma}^{(n)}$ can be numerically computed for any channel by Monte Carlo simulations.

4. Simulation Results

In this section, we show simulation results for the ITU Vehicular-A channel with 6 taps of [7], with a speed of the subscriber station equal to 30 Km/h and the Clarke's model for the Doppler spectrum (see [8], pp. 38–40). In particular, we have considered a (4x4) MIMO channel, with uncorrelated antennas both at the TX side and at the RX side. We have simulated the multilayer schemes described below.

- The 2-layer QPSK has been compared with the single-layer 16QAM, both with the 16e convolutional turbo code (CTC) with rate 1/2. The information word length is 24 bytes for the 2-layer and 48 bytes for the single-layer, yielding the same throughput in the two cases. The resulting spectral efficiency is 8 bits/sec/Hz.

- The 3-layer QPSK has been compared with the single-layer 64-QAM, both with the 16e CTC with rate 2/3. The information word length is 24 bytes for the 2-layer and 48 bytes for the single-layer. Per each layer, 4 codewords are transmitted in each burst for the 2-layer, while 6 codewords are transmitted for the single-layer, yielding the same throughput in the two cases. The resulting spectral efficiency is 16 bits/sec/Hz.

The chosen simulation parameters are compliant with the IEEE 802.16e standard ([2]) and with the simulation scenario guidelines of [7].

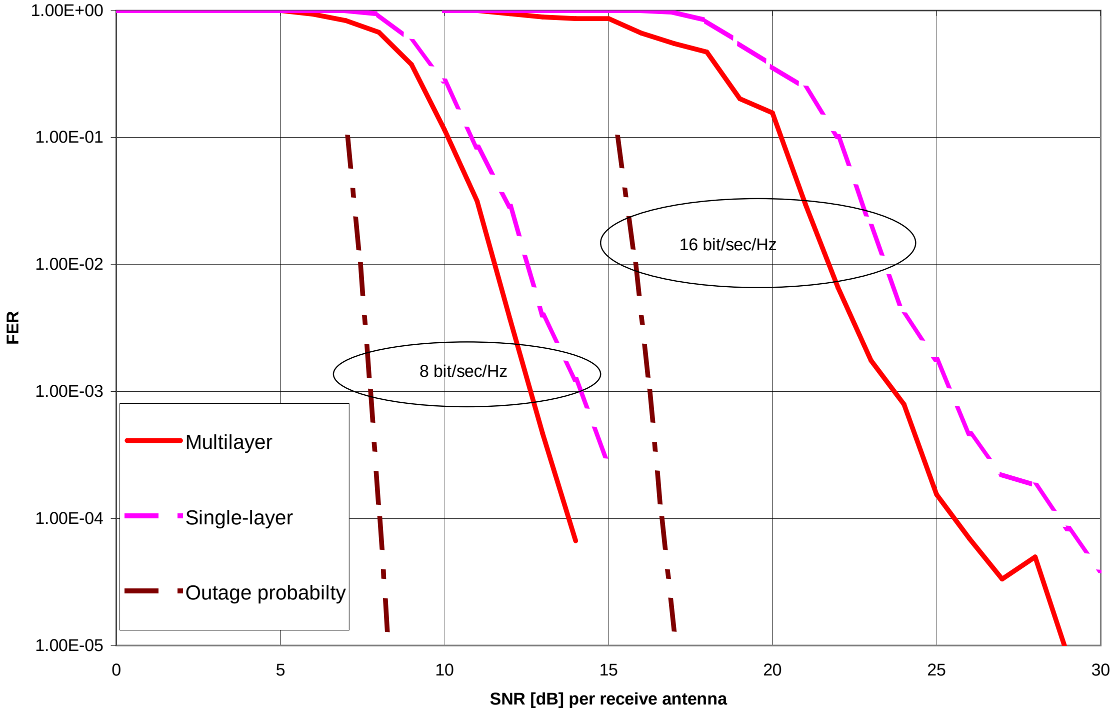

Figure 3: Simulation results. Multilayer schemes: solid lines. Single-layer schemes: dashed lines. Outage probability: dash-dotted lines

The OFDMA subcarrier allocation is based on the uplink PUSC [2], using three adjacent OFDM symbols to transmit. The cyclic prefix length is chosen to be 128 bits. Perfect CSI and perfect power control on OFDMA symbol basis is supposed. Clearly, these assumptions are impractical, but they have the advantage of giving the "ideal" performance of the scheme, without the suboptimalities of channel estimation and power adaptation algorithms. At the receiver, we allow a maximum of 10 iterations both for the outer loop (of the overall iterative receiver) and, per outer iteration, for the inner loop (of the CTC turbo decoder).

In Fig. 3, we show the simulation results in terms of FER. On the abscissa, we show the channel SNR per receive antenna. As it can be seen, the multilayer schemes are always outperforming the single-layer ones with the same spectral efficiency by about 1dB. We also plot the outage probability curves, which show a gap of about 3-4 dB for the 2-layer scheme and of about 8 dB for the 3-layer scheme.

5. Conclusions

In this paper, we have introduced a multilayer spatial multiplexing scheme, iteratively (turbo) decoded at the receiver, which has the following properties:

- it is very flexible, because it is fit for any number of transmit antennas and allows decoupling the number of layers (K) with the multiplexing gain (which, in our paper, was equal to the number of transmit antennas, N_t);
- it can be optimized w.r.t. the channel statistics and code properties, thanks to the degrees of freedom of the combination coefficients; the semi-analytical technique described in Sect. 3., although suboptimal, is simple and gives close-to-optimum results, at least as we have observed in the simulated scenarios;

- it allows for an iterative (turbo-like) decoder, which shows a very good performance for high spectral efficiencies, at an affordable computational complexity;
- it gains 1 dB over single-layer schemes with the same spectral efficiency, while having a lower complexity than the latter. This is because, in the multilayer case, the front-end suboptimally demodulates antenna symbols by considering them as the superposition of K QPSK symbols, while, in the single-layer case, the front-end optimally demodulates the antenna symbols.

In the future work, we will investigate:

- the performance of the same structure, when the SISO front-end is substituted with other front-ends, like the one in [5], with a smaller computational complexity;
- how to reduce the gap from the outage probability, which is still quite high, as shown in Fig. 3;
- what is the impact of a channel estimation algorithm and of an imperfect power control on the performance of the iterative receiver.

Acknowledgment

The research leading to these results has received funding from the European Community's Seventh Framework Programme under grant agreement n 215167, project WiMAGIC.

References

[1] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.

[2] "Ieee 802.16e-2005: Ieee standard for local and metropolitan area networks," Feb 2006.

[3] K. Wu and L. Ping, "A quasi-random approach to spacetime codes," *IEEE Transactions on Information Theory*, vol. 54, pp. 1073–1085, March 2008.

[4] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded cdma," *IEEE Transactions on Communications*, vol. 47, pp. 1046–1061, July 1999.

[5] A. Tarable, G. Montorsi, and S. Benedetto, "A linear front end for iterative soft interference cancellation and decoding in coded cdma," *IEEE Wireless Communications*, vol. 4, pp. 507–518, March 2005.

[6] A. Tarable, G. Montorsi, and S. Benedetto, "A suboptimum iterative decoder for space-time trellis codes," in *Proc. IEEE Int. Conf. Commun.*, (Paris, France), pp. 3000–3004, June 2004.

[7] "Ieee 802.16m evaluation methodology document (emd)," July 2008.

[8] D. Tse and P. Viswanath, *Fundamentals of wireless communications*. Cambridge: Cambridge University Press, 2005.