EFFICIENT SOFT DEMODULATION IN MIMO-OFDM SYSTEMS WITH BICM AND CONSTANT MODULUS ALPHABETS

Dominik Seethaler, Gerald Matz, and Franz Hlawatsch

Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology
Gusshausstrasse 25/389, A-1040 Vienna, Austria (Europe)
phone: +43 1 58801 38958, fax: +43 1 58801 38999, email: dominik.seethaler@tuwien.ac.at
web: http://www.nt.tuwien.ac.at

ABSTRACT

MIMO-OFDM using *bit-interleaved coded modulation* (BICM) is an attractive scheme for wireless communications in frequency-selective fading channels. The BICM decoder requires log-likelihood ratios (LLRs) whose exact computation is extremely costly. In this paper, we present a novel method termed *soft sphere-projection algorithm* (SSPA) that provides approximate LLRs for the case of constant modulus symbol alphabets. The SSPA has very low computational complexity and performs nearly as well as the list sphere decoder. These properties are demonstrated by numerical simulations using synthetic and measured MIMO channels.

1. INTRODUCTION

MIMO-OFDM with *bit-interleaved coded modulation* (BICM) is a promising technique for wireless communications over frequency-selective fading channels (e.g., [1–4]). Such systems usually employ a two-stage receiver consisting of a soft demodulator and a channel decoder. The soft demodulator computes log-likelihood ratios (LLRs)—i.e., soft values—for the coded bits; these LLRs are used by the decoder as bit metrics. Unfortunately, exact LLR calculation is computationally very costly. Various efficient MIMO-BICM soft demodulation algorithms providing approximate LLRs have therefore been proposed, such as the list extension of the Fincke-Phost sphere decoding (LFPSD) algorithm [5] and algorithms based on zero-forcing (ZF) or minimum mean-square error (MMSE) equalization [4,6,7].

In this paper, we present a novel soft demodulation method, termed *soft sphere-projection algorithm* (SSPA), that is applicable to constant modulus alphabets. The SSPA is based on the (hard-decision) sphere-projection algorithm (SPA) [8, 9] and achieves nearly the performance of the LFPSD at a small fraction of the LFPSD's computational complexity. We note that a similar approach using a different hard-detector was described in [10]. We also present methods that exploit subcarrier correlation for a further reduction of the computational complexity of the SSPA in MIMO-OFDM/BICM systems.

The paper is organized as follows. In Section 2, we provide some background on MIMO-OFDM using BICM. The SSPA is outlined in Section 3 and described in more detail in Section 5 (Section 4 reviews the SPA). Simulation results and conclusions are provided in Section 6 and 7, respectively.

2. MIMO-OFDM WITH BICM

System Model. We consider a BICM-based MIMO-OFDM system (see e.g. [2, 4]) with M_T transmit antennas, $M_R \geq M_T$ receive antennas, and K subcarriers. A block of information bits is passed through a channel coder and an interleaver. The coded and interleaved bits are mapped to complex data symbols $d_{m,k} \in \mathscr{A}$ that are transmitted on the mth antenna and kth subcarrier ($m = 1, \ldots, M_T$, $k = 0, \ldots, K-1$). Throughout this paper, we assume Gray labeling and a constant modulus symbol alphabet $\mathscr{A}$ with $|d_{m,k}| = 1$.

The impulse response of the frequency-selective MIMO channel is assumed to have L taps $\mathbf{H}[l]$, $l = 0, \ldots, L-1$. The MIMO-OFDM modulator and demodulator convert the frequency-selective MIMO channel into K parallel flat fading MIMO channels acting as

$$\mathbf{r}_k = \mathbf{H}_k \mathbf{d}_k + \mathbf{w}_k, \qquad k = 0, \ldots, K-1. \tag{1}$$

Here, $\mathbf{r}_k$ denotes the received vector, $\mathbf{H}_k = \sum_{l=0}^{L-1} \mathbf{H}[l] \, e^{-j2\pi \frac{kl}{K}}$ is the channel matrix, $\mathbf{d}_k \triangleq (d_{1,k} \cdots d_{M_T,k})^T \in \mathscr{D} \triangleq \mathscr{A}^{M_T}$ is the transmit vector, and $\mathbf{w}_k$ is additive white Gaussian noise.

Using $\mathbf{r}_k$ and $\mathbf{H}_k$, the soft demodulator calculates (approximate) LLRs for each of the $M_T \log_2 |\mathscr{A}|$ code bits associated with $\mathbf{d}_k$. After deinterleaving, the LLRs from all layers and subcarriers serve as bit metrics for the soft-input channel decoder.

LLR Calculation. In what follows, we will omit the subcarrier index k to simplify notation. Let $b_{m,i}$, $i = 1, \ldots, \log_2 |\mathscr{A}|$ denote the bit label of the symbol $d_m \in \mathscr{A}$ of the mth layer. We assume that the code bits $b_{m,i}$ are i.i.d. uniform. Using the log-sum approximation [2, 5] and ignoring scaling factors irrelevant to our setup, the LLR for $b_{m,i}$ equals

$$\Lambda_{m,i} \triangleq \lambda_{m,i}^0 - \lambda_{m,i}^1, \qquad \text{with} \quad \lambda_{m,i}^b \triangleq \min_{\mathbf{d} \in \mathscr{D}_{m,i}^b} \psi^2(\mathbf{d}). \tag{2}$$

Here, $\psi^2(\mathbf{d}) \triangleq \|\mathbf{r} - \mathbf{H}\mathbf{d}\|^2$ and $\mathscr{D}_{m,i}^b$ denotes the set of all transmit vectors whose label at layer m and bit position i equals $b \in \{0, 1\}$. Note that $\mathscr{D}_{m,i}^b$ consists of all data vectors $\mathbf{d}$ with $d_m \in \mathscr{A}_i^b$ and $d_{m'} \in \mathscr{A}$, $m' \neq m$, where the reduced alphabet $\mathscr{A}_i^b$ is obtained by retaining only those symbols of $\mathscr{A}$ whose bit label at position i equals b. In practice, exact soft demodulation according to (2) may be unfeasible since its complexity scales exponentially with M_T.

Soft demodulation is closely related to the ML detector

$$\hat{\mathbf{d}}_{\mathrm{ML}} = \underset{\mathbf{d} \in \mathscr{D}}{\arg\min} \|\mathbf{r} - \mathbf{H}\mathbf{d}\|^2 = \underset{\mathbf{d} \in \mathscr{D}}{\arg\min} \, \psi^2(\mathbf{d}). \tag{3}$$

Denoting the ith bit of the mth component of $\hat{\mathbf{d}}_{\mathrm{ML}}$ as $(\hat{\mathbf{d}}_{\mathrm{ML}})_{m,i} = b$, it follows that $\hat{\mathbf{d}}_{\mathrm{ML}} \in \mathscr{D}_{m,i}^b$ and hence we immediately obtain

$$\lambda_{m,i}^b = \psi^2(\hat{\mathbf{d}}_{\mathrm{ML}}). \tag{4}$$

However, (3) also has exponential complexity and one must still compute the term $\lambda_{m,i}^{\bar{b}}$ in (2) ($\bar{b} = 1 - b$ denotes bit flipping).

3. BASIC IDEA OF THE SSPA

The idea behind the SSPA is as follows. The SPA [8,9] is an efficient approximation to the ML detector. It replaces $\mathscr{D}$ with a *reduced search set* $\tilde{\mathscr{D}}$ to obtain a hard decision

$$\hat{\mathbf{d}}_{\text{SP}} \triangleq \arg\min_{\mathbf{d}\in\tilde{\mathscr{D}}} \psi^2(\mathbf{d}) \qquad (5)$$

such that $\psi^2(\hat{\mathbf{d}}_{\text{SP}}) \approx \psi^2(\hat{\mathbf{d}}_{\text{ML}})$ (Section 4 will provide more details). This together with (4) yields the approximation

$$\lambda_{m,i}^b \approx \tilde{\lambda}_{m,i}^b \triangleq \psi^2(\hat{\mathbf{d}}_{\text{SP}}), \qquad \text{for} \quad b = (\hat{\mathbf{d}}_{\text{SP}})_{m,i}.$$

It remains to calculate a similar approximation for $\lambda_{m,i}^{\bar{b}}$. A comparison of (2) and (4) reveals that such an approximation could in principle be obtained by applying the SPA with $\mathscr{D}_{m,i}^{\bar{b}}$ instead of $\mathscr{D}$, i.e.,

$$\lambda_{m,i}^{\bar{b}} \approx \tilde{\lambda}_{m,i}^{\bar{b}} \triangleq \min_{\mathbf{d}\in\tilde{\mathscr{D}}_{m,i}^{\bar{b}}} \psi^2(\mathbf{d}). \qquad (6)$$

Here $\tilde{\mathscr{D}}_{m,i}^{\bar{b}}$ is the reduced search set corresponding to $\mathscr{D}_{m,i}^{\bar{b}}$. In general, computing (6) requires a second SPA pass. However, we will show in Section 5 that this second SPA pass can be circumvented and $\tilde{\lambda}_{m,i}^{\bar{b}}$ can be calculated with little additional effort using intermediate results of the first SPA pass. In fact, all vectors $\tilde{\mathbf{d}} \in \tilde{\mathscr{D}}_{m,i}^{\bar{b}}$ can be obtained by a *requantization* of the vectors $\mathbf{d} \in \tilde{\mathscr{D}}$ and the associated distances $\psi^2(\tilde{\mathbf{d}})$ can be computed via simple updates of $\psi^2(\mathbf{d})$ (see Section 5 and in particular (9)). The SSPA for one subcarrier can be summarized as follows:

Step 1: Use the SPA to calculate the reduced search set $\tilde{\mathscr{D}}$, the associated distance set $\Psi = \{\psi^2(\mathbf{d}), \mathbf{d} \in \tilde{\mathscr{D}}\}$, the detector output $\hat{\mathbf{d}}_{\text{SP}}$, and $\psi^2(\hat{\mathbf{d}}_{\text{SP}})$, the minimum element of Ψ.

Step 2: For each $m \in \{1,\dots,M_{\text{T}}\}$ and $i \in \{1,\dots,\log_2|\mathscr{A}|\}$:

- determine $b = (\hat{\mathbf{d}}_{\text{SP}})_{m,i}$ and set $\tilde{\lambda}_{m,i}^b = \psi^2(\hat{\mathbf{d}}_{\text{SP}})$;

- calculate $\Psi_{m,i}^{\bar{b}} = \{\psi^2(\mathbf{d}), \mathbf{d} \in \tilde{\mathscr{D}}_{m,i}^{\bar{b}}\}$ by requantizing each $\mathbf{d} \in \tilde{\mathscr{D}}$ and performing the distance update (9);

- obtain $\tilde{\lambda}_{m,i}^{\bar{b}}$ as the minimum element of $\Psi_{m,i}^{\bar{b}}$;

- finally, calculate the approximate LLR

 $$\Lambda_{m,i} \approx (1-2b)\big(\tilde{\lambda}_{m,i}^b - \tilde{\lambda}_{m,i}^{\bar{b}}\big),$$

  where the factor $1-2b \in \{-1,1\}$ serves to adjust the sign.

We recall that the SSPA has to be performed for each subcarrier separately (cf. (1)). For $M_{\text{T}} = M_{\text{R}}$, the overall complexity can be shown to scale as $\mathcal{O}(KM_{\text{T}}^3|\mathscr{A}|\log_2|\mathscr{A}|)$.

4. REVIEW OF THE SPA

Before explaining the SSPA in more detail, we provide a short review of the hard-decision SPA (see [8,9] for further details). In essence, the SPA is an add-on to an arbitrary reference detector (in this paper we restrict to ZF and MMSE detectors) that uses a reduced search set $\tilde{\mathscr{D}}$ to approximate the ML detector (cf. (5)). The reduced search set is given by $\tilde{\mathscr{D}} \triangleq \{\hat{\mathbf{d}}\} \cup \mathscr{P}$. Here, $\hat{\mathbf{d}} = Q\{\mathbf{y}\}$ is the reference detector output ($\mathbf{y}$ is the output of the ZF or MMSE equalizer and $Q\{\mathbf{y}\}$ denotes elementwise quantization with respect to $\mathscr{A}$) and $\mathscr{P}$ is an additional search set (see next paragraph). Apart from $\tilde{\mathscr{D}}$, the SPA provides the corresponding distance set $\Psi = \{\psi^2(\mathbf{d}), \mathbf{d} \in \tilde{\mathscr{D}}\}$ and $\hat{\mathbf{d}}_{\text{SP}} = \arg\min_{\mathbf{d}\in\tilde{\mathscr{D}}} \psi^2(\mathbf{d})$. One can show that $|\tilde{\mathscr{D}}| \leq |\mathscr{A}|M_{\text{T}}+2$ (e.g., for $M_{\text{T}} = 6$ and 4-QAM, the size of the reduced search set $\tilde{\mathscr{D}}$ is less than 26 while the full search set $\mathscr{D}$ has
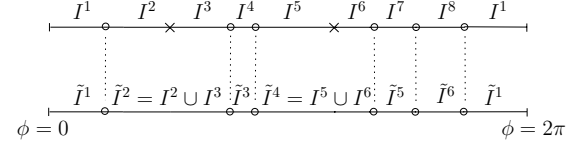

(a) Component circles:
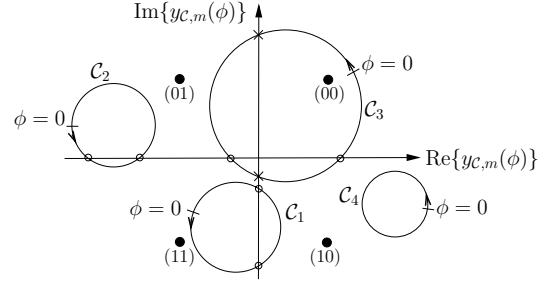
(b) Angle intervals:

Figure 1: *Illustration of SPA geometry for a 4×4 MIMO channel and 4-QAM modulation: (a) component circles $\mathscr{C}_m$ and (b) associated angle intervals I^r (corresponding to $\mathscr{P}$) and $\tilde{I}^r$ (corresponding to $\mathscr{P}_{3,2}^0$). Here, $\mathscr{A}_2^0 = \left\{\frac{1+j}{\sqrt{2}}, \frac{1-j}{\sqrt{2}}\right\}$.*

size 4096). The computational complexity of the SPA (not counting the complexity of the reference detector) is $\mathcal{O}(|\mathscr{A}|M_{\text{T}}^2)$.

The construction of $\mathscr{P}$ is based on approximating $\mathbf{H}$ by an *idealized bad channel* (IBC). Furthermore, it exploits the constant modulus property which implies that all data vectors lie on an M_{T}-dimensional hypersphere $\mathscr{H}$, i.e., $\mathscr{D} \subset \mathscr{H}$. The center of this hypersphere is the origin and its radius equals $R = \sqrt{M_{\text{T}}}$. The IBC captures essential properties of a poorly conditioned channel $\mathbf{H}$ by approximating the smallest singular value of $\mathbf{H}$ with zero and the remaining singular values with the largest singular value.

In what follows, let $\mathbf{v}$ denote the right singular vector of $\mathbf{H}$ corresponding to the smallest singular value. With the IBC approximation, it is shown in [8,9] that $\psi^2(\mathbf{d})$ is small if and only if $\mathbf{d}$ is close to the reference line

$$\mathscr{L}: \quad \mathbf{y}_{\mathscr{L}}(\alpha) = \alpha\mathbf{v} + \mathbf{y}^\perp, \qquad \alpha \in \mathbb{C}, \qquad (7)$$

where $\mathbf{y}^\perp = (\mathbf{I} - \mathbf{v}\mathbf{v}^H)\mathbf{y}$ denotes the component of $\mathbf{y}$ orthogonal to $\mathbf{v}$. For that reason, $\mathscr{P}$ will be composed of data vectors that are close to $\mathscr{L}$. Here, two cases can be distinguished.

Case I. If $\mathscr{L}$ intersects $\mathscr{H}$, $\mathscr{P}$ is chosen to consist of data vectors close to the intersection circle $\mathscr{C} = \mathscr{L} \cap \mathscr{H}$ that is given by

$$\mathscr{C}: \quad \mathbf{y}_{\mathscr{C}}(\phi) \triangleq \sqrt{R^2 - \|\mathbf{y}^\perp\|^2}\, e^{j\phi}\mathbf{v} + \mathbf{y}^\perp, \qquad \phi \in [0, 2\pi).$$

Evidently, the data vector closest to $\mathbf{y}_{\mathscr{C}}(\phi)$ equals $Q\{\mathbf{y}_{\mathscr{C}}(\phi)\}$. Thus, $\mathscr{P}$ consists of all data vectors obtained as $Q\{\mathbf{y}_{\mathscr{C}}(\phi)\}$ for all possible angles ϕ. For an efficient construction of $\mathscr{P}$, $\mathscr{C}$ is viewed as the collection of M_{T} component circles $\mathscr{C}_m$ in the symbol domain (see Fig. 1(a)). If we move along $\mathscr{C}$ by varying ϕ, the mth component of $\mathbf{y}_{\mathscr{C}}(\phi)$ moves along $\mathscr{C}_m$. The first data vector in $\mathscr{P}$ is obtained as $\mathbf{d}^1 = Q\{\mathbf{y}_{\mathscr{C}}(0)\}$. Then ϕ is increased until one of the components of $\mathbf{y}_{\mathscr{C}}(\phi)$ crosses the boundary of a symbol decision region; at that point, a second data vector $\mathbf{d}^2$ is added to $\mathscr{P}$. This procedure is continued until we have moved along the whole intersection circle $\mathscr{C}$. Thus, each data vector $\mathbf{d}^r \in \mathscr{P}$, $r = 1,\dots,|\mathscr{P}|$ is associated with a specific interval I^r of the angle parameter ϕ, i.e., $\mathbf{d}^r = Q\{\mathbf{y}_{\mathscr{C}}(\phi)\}$ for any $\phi \in I^r$ (cf. Fig. 1).

Case II. If $\mathscr{L}$ and $\mathscr{H}$ do not intersect, then $\mathbf{y}^{\perp}$ is the point of $\mathscr{L}$ closest to $\mathscr{H}$ [8], and thus $\mathbf{d}^1 = Q\{\mathbf{y}^{\perp}\}$. The remaining data vectors in $\mathscr{P}$ are chosen as the nearest neighbors of $\mathbf{d}^1$, which can be easily found by substituting the two nearest-neighbor symbols for the individual components of $\mathbf{d}^1$.

5. SSPA IMPLEMENTATION DETAILS

Preliminaries. Recall that $\mathscr{D}_{m,i}^{\bar{b}}$ consists of all data vectors $\mathbf{d}$ with $d_m \in \mathscr{A}_i^{\bar{b}}$ and $d_{m'} \in \mathscr{A}$, $m' \neq m$. Thus, when determining the reduced search set $\tilde{\mathscr{D}}_{m,i}^{\bar{b}} = \{\hat{\mathbf{d}}_{m,i}^{\bar{b}}\} \cup \mathscr{P}_{m,i}^{\bar{b}}$ for (6), the main difference to the initial SPA pass is that quantization and nearest neighbor search for layer m are limited to the reduced alphabet $\mathscr{A}_i^{\bar{b}}$. The efficient implementation of the SSPA is based on the following fundamental observation. For Gray labeling and a constant modulus alphabet whose size is a power of two, it can be shown that quantization with respect to the reduced alphabet $\mathscr{A}_i^{\bar{b}}$ (denoted $Q_i^{\bar{b}}\{\cdot\}$) remains unchanged when preceded by quantization with respect to $\mathscr{A}$, i.e.,

$$Q_i^{\bar{b}}\{y\} = Q_i^{\bar{b}}\{Q\{y\}\}. \tag{8}$$

This means that the decision boundaries for $\mathscr{A}_i^{\bar{b}}$ form a subset of the boundaries for $\mathscr{A}$. Thus, any quantization with respect to $\mathscr{A}_i^{\bar{b}}$ can be performed using $\mathscr{A}$-quantized values. This allows one to determine $\tilde{\mathscr{D}}_{m,i}^{\bar{b}}$ solely by using $\tilde{\mathscr{D}}$ (see below).

The basic operation to obtain a data vector $\tilde{\mathbf{d}} \in \tilde{\mathscr{D}}_{m,i}^{\bar{b}}$ is the *requantization* of the mth component of $\mathbf{d} \in \tilde{\mathscr{D}}$ (the other components remain unchanged):

$$\tilde{d}_{m'} = \begin{cases} Q_i^{\bar{b}}\{d_m\}, & \text{if } m' = m, \\ d_{m'}, & \text{if } m' \neq m. \end{cases}$$

On the vector level, this requantization of the mth component will be denoted $\tilde{\mathbf{d}} = Q_{m,i}^{\bar{b}}\{\mathbf{d}\}$. For the reference detector output, we thus have $\hat{\mathbf{d}}_{m,i}^{\bar{b}} = Q_{m,i}^{\bar{b}}\{\hat{\mathbf{d}}\}$. It remains to determine $\mathscr{P}_{m,i}^{\bar{b}}$.

Construction of $\mathscr{P}_{m,i}^{\bar{b}}$. If $\mathscr{L}$ intersects $\mathscr{H}$ (case I in Section 4), each data vector $\mathbf{d}^r \in \mathscr{P}$ is associated to a specific interval I^r, i.e., $\mathbf{d}^r = Q\{\mathbf{y}_{\mathscr{C}}(\phi)\}$ for any $\phi \in I^r$. For $\mathscr{P}_{m,i}^{\bar{b}}$, we now use $\mathscr{A}_i^{\bar{b}}$ instead of $\mathscr{A}$ for layer m. Thus, each data vector $\tilde{\mathbf{d}}^r \in \mathscr{P}_{m,i}^{\bar{b}}$ is obtained as $\tilde{d}_m^r = Q_i^{\bar{b}}\{y_{\mathscr{C},m}(\phi)\}$ and $\tilde{d}_{m'}^r = Q\{y_{\mathscr{C},m'}(\phi)\}$, $m' \neq m$, for ϕ in a specific interval $\tilde{I}^r$. Property (8) implies that an interval $\tilde{I}^r$ either equals an interval I^r or is a union of adjacent intervals I^r (cf. Fig. 1(b)). This is due to the fact that some decision boundaries for $\mathscr{A}$ and the corresponding intersections with the mth component circle $\mathscr{C}_m$ are no longer present for the reduced alphabet $\mathscr{A}_i^{\bar{b}}$. It follows that the *entire* set $\mathscr{P}_{m,i}^{\bar{b}}$ can be obtained by requantizing every data vector $\mathbf{d}^r \in \mathscr{P}$, i.e., $\tilde{\mathbf{d}}^r = Q_{m,i}^{\bar{b}}\{\mathbf{d}^r\}$. In that way, data vectors $\tilde{\mathbf{d}}^r$ associated to an interval $\tilde{I}^r$ that is the union of two or more intervals I^r will be multiply obtained. We do not take this into account, since the resulting complexity reduction would be insignificant.

If $\mathscr{L}$ and $\mathscr{H}$ do not intersect (case II in Section 4), $\mathscr{P}$ consists of $\mathbf{d}^1 = Q\{\mathbf{y}^{\perp}\}$ and all its nearest neighbors. Thus, $\tilde{\mathbf{d}}^1 \in \mathscr{P}_{m,i}^{\bar{b}}$ is obtained as $\tilde{\mathbf{d}}^1 = Q_{m,i}^{\bar{b}}\{\mathbf{d}^1\}$. Furthermore, for constant modulus alphabets it can be checked quite easily that $Q_{m,i}^{\bar{b}}\{\mathbf{d}^r\}$, $\mathbf{d}^r \in \mathscr{P}$, $r = 2, \ldots, |\mathscr{P}|$ yields the *entire* set of nearest neighbors of $\tilde{\mathbf{d}}^1$.

Efficient Calculation of $\Psi_{m,i}^{\bar{b}}$. To calculate $\tilde{\lambda}_{m,i}^{\bar{b}}$ according to (6), we need the distance set $\Psi_{m,i}^{\bar{b}}$ associated to $\tilde{\mathscr{D}}_{m,i}^{\bar{b}}$. The evaluation of $\psi^2(\tilde{\mathbf{d}})$ has roughly complexity $\mathscr{O}(M_{\text{T}}M_{\text{R}})$. We next show that the distances $\psi^2(\tilde{\mathbf{d}})$, $\tilde{\mathbf{d}} \in \tilde{\mathscr{D}}_{m,i}^{\bar{b}}$ can be obtained via a low-complexity update of $\psi^2(\mathbf{d})$, $\mathbf{d} \in \tilde{\mathscr{D}}$, provided that the error vectors $\mathbf{e}(\mathbf{d}) \triangleq \mathbf{r} - \mathbf{H}\mathbf{d}$ were stored during the first SPA pass. Consider the requantized data vector $\tilde{\mathbf{d}} = Q_{m,i}^{\bar{b}}\{\mathbf{d}\}$ and let $\Delta_{m,i}^{\bar{b}} \triangleq \tilde{d}_m - d_m$ denote the difference between the mth components of $\tilde{\mathbf{d}}$ and $\mathbf{d}$. It can then be shown that

$$\psi^2(\tilde{\mathbf{d}}) = \psi^2(\mathbf{d}) + \|\mathbf{h}_m\|^2 |\Delta_{m,i}^{\bar{b}}|^2 - 2\operatorname{Re}\left\{\mathbf{e}^H(\mathbf{d})\,\mathbf{h}_m\,\Delta_{m,i}^{\bar{b}}\right\}, \tag{9}$$

where $\mathbf{h}_m$ denotes the mth column of $\mathbf{H}$. This update requires roughly $\mathscr{O}(M_{\text{R}})$ operations.

Exploiting Subcarrier Correlation. In a MIMO-OFDM system with BICM, the SSPA has to be run for each subcarrier. Further computational savings are possible by exploiting the strong correlation of the MIMO channel matrices $\mathbf{H}_k$ of nearby subcarriers [11] (now, we take the subcarrier index k again into account, see (1)). In particular, this allows us to use efficient interpolation schemes (e.g. [11]) to calculate the inverses of $\mathbf{H}_k$, $k = 1, \ldots, K$ required by the ZF or MMSE reference detector.

For the SSPA, we further need the right singular vector $\mathbf{v}_k$ corresponding to the smallest singular value of $\mathbf{H}_k$ (see (7)). Since $\mathbf{v}_k$ also is the eigenvector of the inverse Gram matrix[1] $\mathbf{G}_k = (\mathbf{H}_k^H \mathbf{H}_k)^{-1}$ associated to the largest singular value, it can be efficiently computed by means of a few iterations of the power method [12]:

$$\mathbf{v}_k^{(j)} = \frac{\mathbf{G}_k \mathbf{v}_k^{(j-1)}}{\left\|\mathbf{G}_k \mathbf{v}_k^{(j-1)}\right\|}, \qquad j = 1, \ldots, J.$$

Usually, these iterations are initialized with a random vector $\mathbf{v}_k^{(0)}$. However, since it can be expected that $\mathbf{v}_k$ does not change dramatically from one subcarrier to another, we propose to initialize the power method with the singular vector computed for the preceding subcarrier, i.e., $\mathbf{v}_k^{(0)} = \mathbf{v}_{k-1}^{(J)}$. In general, this significantly reduces the number of iterations required. In our numerical simulations (see next section), we even used $J = 1$, i.e., just a *single* iteration (matrix-vector product plus normalization) for the singular vector update (apart from the first subcarrier where we used a random initialization and $J = 5$).

6. SIMULATION RESULTS

We next provide simulation results to illustrate the packet error rate (PER) performance and computational complexity of the SSPA for synthetic and measured frequency-selective MIMO channels. We simulated a BICM-based MIMO-OFDM system with $M_{\text{T}} = M_{\text{R}} = 4$ transmit and receive antennas, $K = 128$ subcarriers, a rate-$1/2$ 16-state convolutional code with octal generators (23, 35), 4 bits trellis termination, a random block interleaver, 4-QAM symbol alphabet, and Gray labeling (thus, each MIMO-OFDM symbol carries 508 information bits). At the receiver, a Viterbi decoder with a traceback depth of 25 was employed for channel decoding. As soft demodulators, we used the SSPA in conjunction with ZF and MMSE detectors as reference detectors (denoted SSPA-Z and SSPA-M, respectively), the LFPSD [5], and ZF-based and MMSE-based soft demodulation according to [6] and [7], respectively. SSPA-Z was considered besides SSPA-M since it has the advantage that knowledge of the noise variance is not required. The LFPSD used LLR thresholds ± 8 and 32 candidate data vectors inside the hypersphere.

---

[1]We note that the $\mathbf{G}_k$'s can also be computed efficiently using the interpolation techniques in [11].
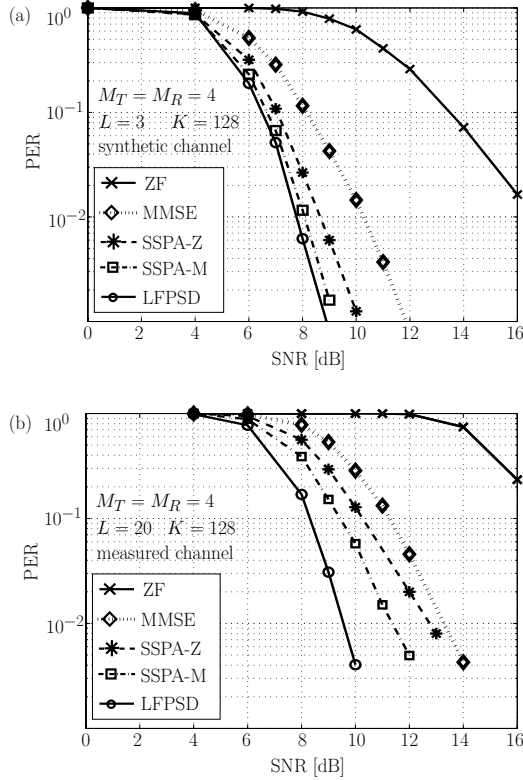
Figure 2: *PER performance of various soft-demodulation schemes versus SNR for (a) a synthetic MIMO channel, (b) a measured MIMO channel.*

Synthetic Channel. We first consider a synthetic MIMO channel with $L = 3$. The channel taps $\mathbf{H}[l]$ are spatially and temporally i.i.d. The PER obtained with the different soft demodulation algorithms versus SNR is shown in Fig. 2(a). For a PER of 10^{-2}, SSPA-M achieves virtually the same performance as LFPSD and SSPA-Z performs less than 1 dB worse. ZF-based and MMSE-based demodulation perform more than, respectively, 8 dB and 2 dB worse than their SSPA counterparts. In contrast to ZF/MMSE-based demodulation, SSPA-M apparently does not incur any diversity loss.

Measured MIMO Channel. Next we evaluate all schemes using indoor MIMO channel measurements obtained at Vienna Airport.[2] The transmitter and receiver positions were fixed and there was no line of sight. The channel data comprised 4452 impulse response snapshots of length $L = 20$. Fig. 2(b) again shows PER versus SNR for the various soft demodulators. While in general PER performance for the measured channel (specifically the coding gain) is somewhat poorer, the relations of the individual algorithms to each other are quite similar to the synthetic case. To achieve a PER of 10^{-2}, an SNR of about 9.5 dB, 11.3 dB, 12.8 dB, 13.3 dB, and 20 dB is required for LFPSD, SSPA-M, SSPA-Z, MMSE, and ZF, respectively. The gap between SSPA-M and LFPSD now is larger, which can be attributed to spatial channel correlations that reduce the quality of the IBC approximation underlying the SSPA.

Computational Complexity. For a rough comparison of the computational complexity of the various demodulators, we provide MATLAB kflops (per subcarrier and OFDM symbol) measured in

---

the synthetic simulation setup within 100 OFDM symbols at an SNR of 7 dB. We did not count the calculations for the equalizers (required for ZF/MMSE and SSPA) and QR-decompositions (required for LFPSD). With LFPSD and SSPA, complexity depends on the individual channel realization. The LFPSD required up to 118.2 kflops with an average of 44 kflops. With the SSPA, on average 4.2 and at most 6.7 kflops were required. The complexity of ZF/MMSE based demodulation was 0.3 kflops. We conclude that SSPA is more complex than ZF and MMSE, but an order of magnitude more efficient than LFPSD.

7. CONCLUSIONS

The *soft sphere-projection algorithm* (SSPA) is a novel soft demodulation technique for MIMO-OFDM systems using bit-interleaved coded modulation (BICM). The SSPA is based on the (hard-decision) sphere-projection algorithm (SPA) and exploits intermediate SPA results to obtain approximate log-likelihood ratios with very low computational effort. Simulation results demonstrated that the performance of the SSPA is similar to that of Fincke-Phost list sphere decoding although the computational complexity is significantly smaller. Thus, the SSPA offers an excellent complexity-performance compromise for MIMO-OFDM/BICM systems.

REFERENCES

[1] J. J. Boutros, F. Boixadera, and C. Lamy, "Bit-interleaved coded modulations for multiple-input multiple-output channels," in *Proc. IEEE Symp. Spread Spectrum Techn. Appl.*, (Parsippany, NJ), pp. 123–126, Sept. 2000.

[2] S. H. Müller-Weinfurtner, "Coding approaches for multiple antenna transmission in fast fading and OFDM," *IEEE Trans. Signal Processing*, vol. 50, pp. 2442–2450, Oct. 2002.

[3] R. Visoz, A. O. Berthet, and J. J. Boutros, "Reduced-complexity iterative decoding and channel estimation for space time BICM over frequency-selective wireless channels," in *Proc. IEEE PIMRC-02*, (Lisbon, Portugal), pp. 1017–1022, Sept. 2002.

[4] K.-B. Song and S. A. Mujtaba, "A low complexity space-frequency BICM MIMO-OFDM system for next-generation WLANs," in *Proc. IEEE Globecom 2003*, (San Francisco, CA), pp. 1059–1063, Dec. 2003.

[5] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Inf. Theory*, vol. 51, pp. 389–399, March 2003.

[6] M. Butler and I. Collings, "A zero-forcing approximate log-likelihood receiver for MIMO bit-interleaved coded modulation," *IEEE Commun. Letters*, vol. 8, pp. 105–107, Feb. 2004.

[7] D. Seethaler, G. Matz, and F. Hlawatsch, "An efficient MMSE-based demodulator for MIMO bit-interleaved coded modulation," in *Proc. IEEE Globecom 2004*, vol. IV, (Dallas, Texas), pp. 2455–2459, Dec. 2004.

[8] H. Artés, D. Seethaler, and F. Hlawatsch, "Efficient detection algorithms for MIMO channels: A geometrical approach to approximate ML detection," *IEEE Trans. Signal Processing*, vol. 51, pp. 2808–2820, Nov. 2003.

[9] D. Seethaler, H. Artés, and F. Hlawatsch, "Efficient near-ML detection for MIMO channels: The sphere-projection algorithm," in *Proc. IEEE Globecom 2003*, vol. IV, (San Francisco, CA), pp. 2089–2093, Dec. 2003.

[10] D. Seethaler, G. Matz, and F. Hlawatsch, "Low-complexity soft demodulation of MIMO-BICM using the line search detector," in *Proc. ICC 2005*, vol. IV, (Seoul, Korea), pp. 2447–2451, May 2005.

[11] M. Borgmann and H. Bölcskei, "Interpolation-based efficient matrix inversion for MIMO-OFDM receivers," in *Proc. Asilomar Conf. Signals, Systems and Computers*, vol. II, (Pacific Grove, CA), pp. 1941–1947, Nov. 2004.

[12] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore: Johns Hopkins University Press, 3rd ed., 1996.

[13] M. Herdin, *Non-Stationary Indoor MIMO Radio Channels*. PhD thesis, Vienna University of Technology, Aug. 2004.

---

[2]The authors thank N. Czink for providing the measurement data. A detailed description of the measurement campaign can be found in [13].