Author(s) Levanen, Toni; Renfors, Markku

Title Improved Energy Efficiency for Wireless SC MIMO Through Data-Dependent Superimposed Training

Citation Levanen, Toni; Renfors, Markku 2012. Improved Energy Efficiency for Wireless SC MIMO Through Data-Dependent Superimposed Training . 5th International Symposium on Communications Control and Signal Processing ISCCSP 2012, Rome, Italy, May 2-4, 2012. International Symposium on Communications Control and Signal Processing Piscataway, NJ, 1-6.

Year 2012

DOI http://dx.doi.org/10.1109/ISCCSP.2012.6217790

Version Post-print

URN http://URN.fi/URN:NBN:fi:tty-201409091424

Improved Energy Efficiency for Wireless SC MIMO Through Data-Dependent Superimposed Training

Toni Levanen and Markku Renfors
Department of Communications Engineering
Tampere University of Technology
P.O.Box 553, FIN-33101, Finland
Email: {firstname.lastname}@tut.fi

Abstract—**In this paper, we compare the energy efficiency of a single carrier multiple-input multiple-output communications in terms of the energy per data bit required to achieve desired performance level or throughput. The comparison on energy efficiency is done between traditional time-domain multiplexed training and a more recently introduced data-dependent superimposed training. We extend our earlier single-input multiple-output system to the multiple-input multiple-output case and show how the data-dependent superimposed training based system can achieve better energy efficiency in small and diversity enabled multiple-input multiple-output links. In addition, we present analytical mean squared error results for DDST based MIMO channel estimation and propose methods to improve their accuracy when using short channel estimators.**

***Keywords*: data-dependent superimposed training, channel estimation, energy efficiency, multiple-input multiple-output communications**

I. Introduction

Multiple-input multiple-output (MIMO) communications provide the possibility to increase the throughput of a wireless link simply by adding transmit-receive (Tx-Rx) antenna pairs. The capacity of a wireless MIMO link increases linearly with the number Tx-Rx antenna pairs (assuming independently fading channels) [1]. The main problem in achieving this capacity with traditional time-domain multiplexed training (TDMT) is the fact that we have to allocate more time slots for the training sequence to enable the estimation of all wireless channels involved in the MIMO communication link. Thus, as each Tx-Rx antenna pair increases capacity, it also increases the amount of training information required to estimate the channel between the new Tx antenna and all of the Rx antennas.

Another approach for channel estimation, that has recently obtained growing interest, is the data-dependent superimposed training (DDST) based channel estimation (see, for example [2] and [3]). In DDST, all the time slots are dedicated for the user data symbols and the training sequence is arithmetically added on top of the user data symbols. The data-dependent portion of the training sequence is typically related to the cyclic mean of the transmitted user data symbols. The removal of the cyclic mean corresponds to removing certain frequency bins from the discrete Fourier-transform (DFT) of the user data symbol sequence. Then, the cyclic pilot sequence is contained in these frequency bins and detected in the receiver without interference from the user data. In the receiver, the missing information can be effectively estimated by iteratively calculating the cyclic mean of the detected symbols [4]. For an interested reader, superimposed pilot based single carrier multiuser/MIMO communications, without the data-dependent component, is studied, for example, in [5] and [6].

In this paper, we compare the energy consumption in terms of average energy per bit over the one sided noise spectral density (E_b/N_0). We show that with DDST, in single carrier multiantenna transmission, one can further improve the energy efficiency in single-input multiple-output (SIMO) scenarios and in MIMO communications. As we increase the number of Tx-Rx pairs or the size of the constellation, the performance gain of the DDST decreases with respect to the TDMT, because the effective SNR of a data symbol with DDST decreases as larger portion of average power is removed (due to increased length of pilot sequence) or the deviation of the cyclic mean component increases with respect to the symbol distances (due to increased size of the constellation).

This paper is organized as follows. In Section II the system model and receiver architecture are introduced as extension of the SIMO model of [4] without any power constraints. The channel estimation and related analysis is presented in Section III. In Section IV we compare the energy efficiency of DDST and TDMT with different antenna configurations. Furthermore, we comment on the performance limiting factors of DDST and try to give an idea of the parameter set in which DDST can overperform TDMT. Finally, in Section V, our conclusion are presented.

II. System Model

The system model originates from a possible future uplink-direction mobile wireless link. Therefore, most of the signal processing is located in the receiver side (in the base station). The conceptual block diagram of the simulated system model is given in Fig. 1.

In the transmitter side, we have traditional signal processing blocks, as channel encoder, channel interleaver, symbol mapper and transmit pulse shape filtering. In the studies considered

This work was supported by the Tampere Graduate School in Information Science and Engineering (TISE), the Nokia Foundation and Academy of Finland (under Project no. 129077, “Hybrid Analog-Digital Signal Processing for Communications Transceivers“).

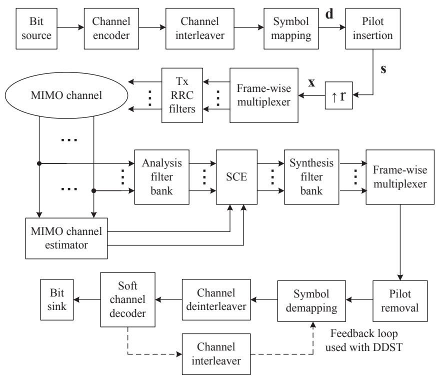

Fig. 1. System model for the single carrier based multiple-input multiple-output simulations.

in this paper, we have concentrated on spatial multiplexing and have not considered any space-time codes. In the receiver side, the maximum number of antennas is limited to four. The MIMO channel frequency domain equalization is done with a filter bank based linear equalizer with MSE criteria, as presented in [7].

The power of the MIMO signal is normalized to unity, thus each Tx antenna transmits with average power equal to $\sigma^2_{s,i} = 1/N_{Tx}$, where i is the index for the transmit antenna. The received signal-to-noise ratio (SNR), based on the given energy per bit over the one sided noise spectral density (E_b/N_0), is given as

$$SNR = \frac{N_b}{rN_s}\frac{E_b}{N_0}, \tag{1}$$

where N_b is the number of user data bits per frame, N_s is the number of symbols per frame and r is the oversampling factor used in the receiver. Note that in every case the number of user data bits is bigger with DDST than with TDMT, and the difference increases as we increase the number of Tx-Rx pairs. This is because we allocate more symbols for training in TDMT that leads to less user data symbols per frame.

Given a certain user data symbol sequence $\mathbf{d}_i$, transmitted from i^{th} Tx antenna, taken from some constellation $\mathbf{d}_i \in \mathcal{C}^N$, where N is the length of the sequence, we can define the symbols after inserting pilots as $\mathbf{s}_i = [\mathbf{p}_i^T \mathbf{d}_i^T]^T$ or $\mathbf{s}_i = \sqrt{(1-\gamma)/(1-1/N_c)}\mathbf{J}\mathbf{d}_i + \sqrt{\gamma}\mathbf{p}_{c,i}$, for TDMT or DDST, respectively. We assume that both, the data sequence $\mathbf{d}_i$ and the cyclic pilot sequence $\mathbf{p}_{c,i}$, have a unit variance $\sigma^2_d = 1$ and $\sigma^2_{p_{c,i}} = 1$, respectively. With DDST, γ defines the fraction of total power allocated to the cyclic pilot sequence $\mathbf{p}_{c,i} = \mathbf{1}_{N_c \times 1} \otimes \mathbf{p}_i$, where $\otimes$ is the Kronecker product and $\mathbf{p}_i$ is a certain cyclic shift of the known basis pilot sequence $\mathbf{p}$, transmitted from the i^{th} Tx antenna. The cyclic mean of the data sequence is removed by multiplying with matrix $\mathbf{J} = \mathbf{I} - 1/N_c \mathbf{1}_{N_c} \otimes \mathbf{I}_{N_p}$. In this paper we assume that the frame length is $N = N_c N_p$, where N_c is the number of cyclic copies and N_p is the length of the basis pilot sequence $\mathbf{p}$. Note, that the basis pilot sequence length is now N_{Tx} times the channel estimator length, allowing us to estimate the channels between all Tx antennas and a single Rx antenna. The signal transmitted from Tx antenna i with DDST is defined as

$$\begin{aligned}\mathbf{x}_i &= \mathbf{H}_{RRC}\sqrt{1/N_{Tx}} \\ &(\sqrt{(1-\gamma)/(1-1/N_c)}\mathbf{J}\mathbf{d}_i + \sqrt{\gamma}\mathbf{p}_{c,i}) \otimes [r\ \underbrace{0\ldots0}_{r-1 \text{ zeros}}],\end{aligned} \tag{2}$$

where $\mathbf{H}_{RRC}$ is a circular matrix containing the Tx pulse shape filter coefficients and is used to calculate the convolution between the Tx pulse shape filter and the symbol sequence. The transmitted symbol sequence is normalized with a factor $1/(1-1/Nc)$, to take into account the average power loss caused by the removal of the cyclic mean component. Kronecker product with vector $\mathbf{r} = [r\ 0\ldots0]$ realizes the r times oversampling with power normalization.

The used basis pilot sequences, $\mathbf{p}$, are so called chirp sequences [8], which are shown to be optimal for DDST in [9] (noted as optimal channel independent (OCI) pilot sequences) and also for TDMT in MIMO communications [10]. The same training is used with both, TDMT and DDST. The only difference is in the ordering of the pilots to each transmitted frame, in order to obtain a full pilot matrix in the receiver for MIMO channel estimation.

In the receiver side, we have channel estimation, filter bank based frequency domain equalization implementing a close-to-optimal linear MIMO detector with heavily frequency-selective channels, pilot removal, symbol demapping and decoding. With DDST, we have a soft iterative loop that is used to estimate the removed cyclic mean of the user data sequence, similar to what was proposed in [4]. The signal received in Rx antenna with index k is defined as

$$\mathbf{y}_k = [\mathbf{H}_{k,1}\ \mathbf{H}_{k,2}\ldots\mathbf{H}_{k,N_{Tx}}][\mathbf{x}_1^T\ \mathbf{x}_2^T\ldots\mathbf{x}_{N_{Tx}}^T]^T + \mathbf{n}_k, \tag{3}$$

where $\mathbf{H}_{k,i}$ represents the convolution matrix of the channel response between i^{th} Tx antenna and k^{th} Rx antenna and $\mathbf{n}_k$ is a vector of complex Gaussian noise components with variance $\sigma^2_n = 1/SNR$. We have normalized the power response of all channel realization to unity.

We have assumed two times oversampling in the receiver $r = 2$, which allows us to efficiently incorporate the RRC filtering in the subchannel wise equalization (SCE) used in the filter bank. The different spatial data streams are obtained by applying the linear subcarrier wise MIMO equalizer to the received signal [7]. The number of subbands in the analysis filter bank is set to 1024 and in the synthesis filter bank it is then 512, due to $r = 2$ times down sampling. The used overlapping factor in the filter bank is equal to 5 and the filter bank roll-off is equal to 1. More details on the filter bank can be found, e.g., in [7], [11] and references therein.

After the channel equalization, we remove the training symbols from the received sequence and normalize the average power per layer to $\sigma^2_{\hat{d}} = 1 + \sigma^2_n$ with TDMT and to $\sigma^2_{\hat{d}} = (1 - 1/N_c) + \sigma^2_n$ with DDST. We have assumed that the noise variance σ^2_n is known by the receiver. With DDST, after the normalization, we add the first cyclic mean estimate to the received symbol sequence. This cyclic mean estimate is based on the hard symbol estimates before channel decoder, as presented in [2]. Then we proceed to the soft channel decoder and with DDST we use maximum of 3, 8, or 10 feedback iterations for soft cyclic mean estimation with 4-, 16-, or 64-QAM modulation, respectively. The soft cyclic mean estimation at each feedback iteration is based on the soft symbol estimates generated from the soft coded bit estimates from the turbo decoder. More details on the soft symbol mapping and demapping can be found, e.g., in [12].

Finally, after all the iterations, we provide the detected bits to the bit sink, from which we eventually obtain bit error rate (BER) and block error rate (BLER) results. We have used BLER to represent goodness of performance of the desired training method and not the frame error rate, because we have defined that each transmitted frame contains q coded blocks, where q is equal to the number of bits per symbol. Thus, each binary coded block in a frame with any constellation is of the same size and corresponds to the number of data symbols per frame. This way the block wise decoding complexity is constant with each constellation and the decoding process can be parallelized in the receiver with larger constellations.

III. Channel estimation and MSE Analysis

In this Section, we introduce the used channel estimator model and derive the related analytic channel estimation MSE. We have used a LS-LMMSE channel estimator, which first uses a least-squares (LS) channel estimator to obtain initial channel estimates and then uses a linear minimum mean-squared-error (LMMSE) channel estimator, based on the LS estimate, to obtain improved channel estimates. The LS channel estimate, for all channels between all Tx antennas and k^{th} Rx antenna is defined as

$$\hat{\mathbf{h}}_{LS,k} = \frac{\mathbf{P}_r^H \sqrt{N_{Tx}}}{r^2 N_p \sigma_p^2} \hat{\mathbf{m}}_{y,k} = \mathbf{h}_k + \frac{\mathbf{P}_r^H \mathbf{H}_{RRC} \hat{\mathbf{m}}_{n,k}}{r^2 N_p \sigma_p^2}, \quad (4)$$

where $\mathbf{h}_k = [\mathbf{h}_{k,1}^T \; \mathbf{h}_{k,2}^T \ldots \mathbf{h}_{k,N_{Tx}}^T]^T$ is the equivalent channel between all Tx antennas and k^{th} Rx antenna, $\mathbf{P}_r^H$ is a cyclic pilot matrix oversampled with factor $r = 2$ and having basis pilot vector $\mathbf{p}$ as its first column vector, $\hat{\mathbf{m}}_{y,k} = \mathbf{J}_{Rx}\mathbf{y}_k$ is the cyclic mean of the received signal vector in antenna k, and the matrix used to calculate the cyclic mean of the received sequences is defined as $\mathbf{J}_{Rx} = 1/N_c \mathbf{1}_{1\times(N_c+1)} \otimes \mathbf{I}_{N_p}$. This matrix is an extended version of the one used in the transmitter, because the received signal is longer in time because of the time dispersive channel. Similarly, $\hat{\mathbf{m}}_{n,k} = \mathbf{J}_{Rx}\mathbf{n}_k$, is the cyclic mean of the received noise component of antenna k. Note, that vector $\hat{\mathbf{h}}_{LS,k} = [\hat{\mathbf{h}}_{LS,k,1}^T \; \hat{\mathbf{h}}_{LS,k,2}^T \ldots \hat{\mathbf{h}}_{LS,k,N_{Tx}}^T]^T$ contains now the channel estimates from all the Tx antennas to the k^{th} Rx antenna. See [10] for further details how the channel estimator is able to obtain channel estimates based on multiple-input single-output (MISO) principle with TDMT. The same ideology holds for DDST and has been used for the results presented in this paper.

In the channel estimator, we approximate the diagonal correlation matrix $\mathbf{C}$ by the instantaneous tap power obtained from the LS channel estimator as

$$\mathbf{C}_{\hat{\mathbf{h}}_{LS,k}} = diag\left\{|\hat{h}_{LS,k}(0)|^2, |\hat{h}_{LS,k}(1)|^2, \cdots, |\hat{h}_{LS,k}(rN_p - 1)|^2\right\}. \quad (5)$$

By assuming the cyclic chirp (OCI) training sequence, the LS-LMMSE estimator can be reduced to

$$\hat{\mathbf{h}}_{LS-LMMSE,k} = \frac{\mathbf{P}_r^H \sqrt{N_{Tx}}}{(\sigma^2 \mathbf{C}_{\hat{\mathbf{h}}_{LS}}^{-1} + r^2 N_p \sigma_p^2 \mathbf{I}_{rN_p \times rN_p})} \hat{\mathbf{m}}_{y,k}. \quad (6)$$

The variable σ^2 corresponds to the total interference power on top of the cyclic mean of the received signal and is given as

$$\sigma^2 = N_{Tx}(1/N_c + 1/N_c^2)\sigma_w^2 \|\mathbf{h}_{RRC}\|^2. \quad (7)$$

Let us define an error vector $\mathbf{e}_{LS,k} = \hat{\mathbf{h}}_{LS,k} - \mathbf{h}_k(\mathbf{\Xi})$, representing the channel estimation error between the estimated channel and a fraction of the true channel corresponding to the estimated part. Because we have not estimated the full equivalent channels, we use indexing set $\mathbf{\Xi}$ to define the samples that are included in the channel estimation process. This idea of short channel estimator was presented in [13], where it was studied with superimposed training. More discussion on the numerical values used is given in Section IV. Now, assuming that the channel taps are i.i.d., the channel estimation mean-squared-error (MSE) for DDST with LS channel estimator is defined as

$$MSE_{LS,k} = trE[\mathbf{e}_{LS,k}\mathbf{e}_{LS,k}^H] = \ldots = \frac{(1/N_c + 1/N_c^2)\sigma_n^2 N_{Tx}}{\gamma r}, \quad (8)$$

and the average MSE over all spatial channels is defined as

$$MSE_{LS} = \frac{1}{N_{Tx}N_{Rx}} \sum_{k=1}^{N_{Rx}} MSE_{LS,k}. \quad (9)$$

In similar manner, for DDST with the LS-LMMSE channel estimator, we can define the channel estimation MSE between all Tx antennas and k^{th} Rx antenna as

$$MSE_{LS-LMMSE,k} = trE[\mathbf{e}_{LS-LMMSE,k}\mathbf{e}_{LS-LMMSE,k}^H] = \ldots = \sum_{l=1}^{N_p} \frac{\sigma_n^4 \sigma_{LS,k}^{-4}(l)}{(\sigma_n^2 \sigma_{LS,k}^{-2}(l) + \gamma r^2 N_p)^2} \sigma_{h,k}^2(l) + \frac{\gamma r^2 N_p N_{Tx}(1 + 1/N_c)\sigma_n^2}{(\sigma_n^2 \sigma_{LS,k}^{-2}(l) + \gamma r^2 N_p)^2}, \quad (10)$$

,where $\sigma_{h,k}^2(l) = E[|h_k(l)|^2]$ is the expected power of the l^{th} tap in equivalent channel $\mathbf{h}_k$ and $\sigma_{LS,k}^2 = \sigma_{h,k}^2(l) + MSE_{LS,k}/N_p$ is the expected power of the l^{th} tap from the LS

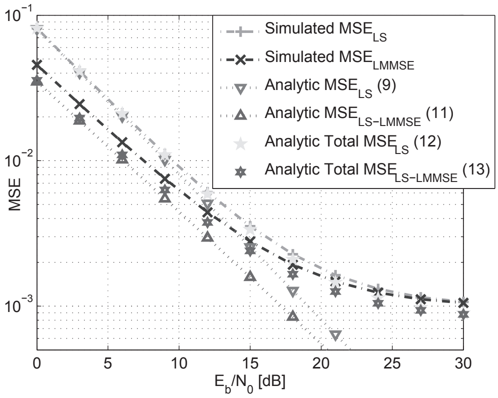

Fig. 2. Channel estimation MSE comparison, for DDST with 2 Tx antennas and 4 Rx antennas, between the simulated MSE and the analytical MSE obtained from (9), (11), (12), and (13).

channel estimator. The average MSE over all spatial channels is defined as

$$MSE_{LS-LMMSE} = \frac{1}{N_{Tx}N_{Rx}} \sum_{k=1}^{N_{Rx}} MSE_{LS-LMMSE,k}. \tag{11}$$

In Fig. 2, we compare the simulated channel estimation MSE with analytical results based on (9) and (11). In this case we have used 2 Tx antennas and 4 Rx antennas. As we can see, there is a error floor in the simulated values which is not present in the analytical ones. This same phenomenon was noted also in [13], which it was proposed that the total MSE of the short channel estimator should include also the expected power of the non-estimated channel taps. Let us now define the analytic MSE estimates based on this idea as

$$Total\ MSE_{LS} = MSE_{LS} + \sum_{s \in \Psi/\Xi} |h_k(s)|^2 \text{and} \tag{12}$$

$$Total\ MSE_{LS-LMMSE} = MSE_{LS-LMMSE} + \sum_{s \in \Psi/\Xi} |h_k(s)|^2, \tag{13}$$

where $\Psi = [0, 1, \ldots, L-1]$ is the indexing set containing all possible index values of the equivalent channel response $\mathbf{h}_k$. The numerical values for our simulations are defined in Section IV.

We have plotted also these MSE results in Fig. 2, and have named them as Total MSE$_{LS}$ and Total MSE$_{LS\text{-}LMMSE}$, for the LS and LS-LMMSE channel estimators, respectively. In [13], the intuition behind the idea is that the channel taps not estimated by the short channel estimator fold on the estimated portion in the cyclic mean computation, and therefore cause the error floor. In our simulations, we have noticed the same phenomenon also with short channel estimation with TDMT, in which the MSE error floor corresponds quite accurately to the expected power of the non-estimated channel taps.

IV. Performance Comparison Between DDST and TDMT

In this section we compare the performance of the SC based MIMO communications with DDST and TDMT in terms of the required average E_b/N_0 to achieve BLER $= 10^{-2}$ or a predefined rate.

The used channel model is the block fading extended ITU Ped. B channel profile [14], which is of length 115 samples with symbol rate $f_{symbol} = 15.36$ MHz and sample rate $f_{sample} = 2f_{symbol}$. In addition, the equivalent channel includes twice the root-raised-cosine (RRC) pulse shape filtering of order $N_{RRC} = 64$, making the equivalent channel length per spatial channel 243 samples long, and the overall channel response observed in the k^{th} receive antenna is of length $L = 243N_{Tx}$ samples. The used RRC filter has a roll-off factor equal to 0.1.

With DDST, we have decided to use channel estimators of length $N_p = (60N_{Tx})$ symbols, if $N_{Tx} = 1, 2,$ or 4, and $N_p = (64N_T x)$ symbols, if $N_{Tx} = 3$. This is to ensure that $N_c = N/N_p$ is an integer. With TDMT the length of the pilot sequence is $N_p = (N_{Tx}+1) * 60 - 1$. This means that we estimate only 120 or 128 samples per spatial channel. This kind of short channel estimator allows us to increase the maximum throughput of the TDMT and to increase the number of cyclic copies with DDST. The channel estimation degradation due to this choice is negligible. As an example, the indexing set for samples estimated with 1,2 or 4 Tx antennas, is then given as

$$\begin{aligned}\Xi \ &= \underbrace{[1\ 1 \ldots 1]}_{N_{Tx}\text{ ones}} \otimes [N_{RRC}/2\ \ N_{RRC}/2+1 \ldots \\ &\quad N_{RRC}/2+119] + 243[0\ 1 \ldots N_{Tx}-1] \otimes \mathbf{1}_{1\times 120}.\end{aligned} \tag{14}$$

In other words, in the channel estimation process, we ignore the first $N_{RRC}/2$ samples and all the samples after index $N_{RRC}/2 + 119$ for each spatial channel.

The used codec in the simulations is a turbo code [15] with generator matrix $G = [1 \frac{1\,5}{1\,3}]$ and the decoding algorithm is the max-log-MAP algorithm with extrinsic information weighting by a factor $\mu = 0.75$ [16]. We have allowed maximum of 5 decoding iterations per code block in the turbo decoder. The channel interleaver and the interleavers inside the turbo codec are S-interleavers [17].

The performance of the DDST based transmission depends on the constellation, number of the cyclic copies of the pilot sequence and on the code rate R. We have done simulations with 4-, 16-, and 64-QAM constellations and with code rates $R = 0.5$, $R = 0.67$, and $R = 0.75$. Let us now discuss on how the different simulation parameters effect on the DDST performance.

As we increase the size of the constellation, it becomes more sensitive to the distortion caused by the removal of the cyclic mean and to the channel estimation errors. Thus, with bigger constellation, higher number of cyclic copies is required to achieve desired level of performance. Increasing pilot power

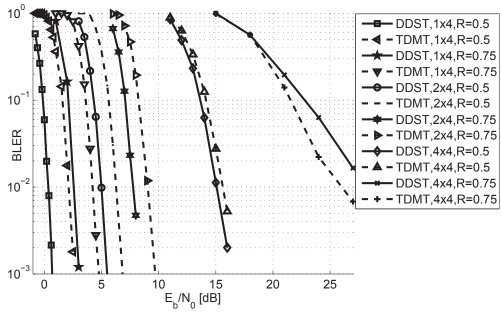

Fig. 3. Block error rate performance with 16-QAM constellation with different MIMO antenna configurations and code rates $R = 0.5$ and $R = 0.75$.

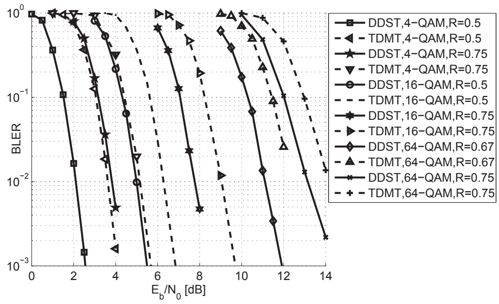

Fig. 4. Block error rate performance with 2x4 MIMO antenna configuration with 4-, 16-, and 64-QAM modulations with code rates $R = 0.5$, $R = 0.67$ and $R = 0.75$.

in this case is a two edged blade, because as we improve the channel estimation MSE, we decrease the effective signal-to-interference-and-noise ratio in the equalizer output. In our simulations, we noticed that the value $\gamma = 0.1$ works rather well in all of the cases. As a rule of thumb, one can consider of allocating the same average power for DDST pilots as is allocated for TDMT pilots, because it was shown in [2], that this leads to the same channel estimation MSE with least-squares type of channel estimators.

The number of cyclic copies depends on the equivalent channel length, channel estimator length and frame length. If the equivalent channel is short with respect to the transmitted frame, we can have higher number of cyclic copies per frame, and thus improve the performance with DDST through improved noise averaging in the channel estimation process and decreased variance of the interference term caused by the cyclic mean removal. If the expected equivalent channel has most of its power concentrated in a relatively short time interval, we can use a short channel estimator that estimates only the significant portion of the channel. With extended ITU Ped. B channel model, we can collect 99.86% of the total power per layer with the used channel estimator of length 120 samples.

Finally, the code rate R affects the performance of the iterative cyclic mean estimation process. The iterative, soft feedback scheme is effective especially with larger constellations and with lower code rates. This is because larger constellations are more sensitive to the cyclic mean estimation errors and smaller code rates allows the turbo decoder to provide more new information for the cyclic mean estimation.

With these short explanations on the expected performance of DDST, we first look at an example of the throughput performance with 16-QAM modulation in different antenna configurations, shown in Fig. 3. In the presented cases the throughput performance of a DDST based system exceeds that of a TDMT based system when we have receiver diversity. In the 4×4 MIMO case DDST is marginally better with coding rate $R = 0.5$, but TDMT provides better performance with $R = 0.75$. One can notice that the difference between DDST and TDMT based systems decreases as we increase the number of Tx-Rx pairs and code rate, and how TDMT eventually exceeds DDST in the 4×4 MIMO case with $R = 0.75$. In 6x6 or 8x8 MIMO antenna configurations TDMT would provide better energy efficiency than DDST, while keeping the other parameters fixed.

In Fig. 4, we compare the throughput performance with different training schemes in 2x4 MIMO antenna configuration. As expected, we can see how the TDMT system starts to achieve the performance of DDST as the constellation size increases or with higher code rates. From the maximal throughput point of view, increasing the number of Tx-Rx pairs and using lower modulation is a more energy efficient and robust method than using constellations larger or equal to 64-QAM.

In Table IV, we have given the best constellation, antenna configuration and code rate combination to achieve the given throughput and the required E_b/N_0 for both, DDST and TDMT. Note that we have simulated discontinuous block fading channel, so these results should be considered as a rough estimate of the required E_b/N_0 to achieve certain throughput in a continuous transmission channel. In all of the cases, the two systems achieve the best E_b/N_0 with the same constellation, code rate and antenna configuration set. We can notice how the TDMT achieves similar E_b/N_0 requirement with high throughput requirement ($\geq$ 150 Mb/s). If we would have more Rx antennas, the DDST based system would be better at even higher throughput rates, given that we could use smaller constellations or provide reception diversity with larger ones. In [6], an iterative receiver for superimposed training was studied, and its performance was compared to coordinated and uncoordinated TDMT users. Also, in [6], the performance of the superimposed training based system was able to compete with the coordinated TDMT based system, if there was reception diversity available. This results is comparable with ours, because we have assumed perfect synchronization for both, DDST and TDMT users. Compared to the superimposed training considered in [6], DDST based system is more sensitive to the frame synchronization, as was shown in [3]. On the other hand, DDST provides better initial

TABLE I
BEST SETUP AND REQUIRED E_b/N_0 FOR DESIRED THROUGHPUT

Throughput	DDST	TDMT
>50 Mb/s	16-QAM, 2x4, R=0.5, E_b/N_0= 4.5 dB	16-QAM, 2x4, R=0.5, E_b/N_0= 6 dB
>100 Mb/s	16-QAM, 3x4, R=0.67, E_b/N_0= 10.5 dB	16-QAM, 3x4, R=0.67, E_b/N_0= 11.5 dB
>150 Mb/s	64-QAM, 3x4, R=0.67, E_b/N_0=18	64-QAM, 3x4, R=0.67, E_b/N_0=18

channel estimates, because the user data does not affect the channel estimation MSE.

V. CONCLUSION

In this paper, we have proposed an analytic channel estimation MSE for DDST in MIMO communications for the LS and LS-LMMSE channel estimator and compared its performance to the simulated channel estimation MSE. In addition, the error floor noted in the simulated MSE results seems to correspond to the sum power of the expected channel taps outside the short channel estimator.

Furthermore, we have compared the required E_b/N_0 with either DDST or TDMT based channel estimation in SC MIMO communications. We have shown, that DDST can overperform the traditional TDMT in small MIMO scenarios (less than 4 Tx-Rx pairs) and when sufficient receiver diversity is available. As typical, there is a complexity penalty to pay for the improved efficiency. Fortunately, the power consumption of the added complexity in the receiver hardware will become smaller than the savings achieved in the transmission power, as the hardware evolves following the famous Moore's law.

The scenario, in which DDST overperforms TDMT, depends on the number of antennas, frame length and on the channel estimator length (or channel length). As long as we have sufficient number of cyclic copies of the basis pilot sequence, we do not cause significant interference on the data symbols and we obtain improved channel estimates through the noise averaging in the DDST based channel estimation. Thus, even though the maximum throughput of the TDMT is significantly smaller in larger MIMO communications (e.g., $k \times k$ MIMO cases, where $k \geq 4$), the interference caused by removing the cyclic mean with DDST may cancel the possible throughput benefits, making TDMT a more viable solution in these cases.

REFERENCES

[1] G. J. Foschini, "Layered Space-Time Architecture for Wireless Communication in a Fading Environment when Using Multi-Element Antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996.
[2] M. Ghogho et al., "Channel Estimation and Symbol Detection for Block Transmission Using Data-Dependent Superimposed Training," *IEEE Signal Process. Lett.*, vol. 12, no. 3, pp. 226–229, 2005.
[3] E. Alameda-Hernandez et al., "Frame/Training Sequence Synchronization and DC-Offset Removal for (Data-Dependent) Superimposed Training Based Channel Estimation," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2557–2569, 2007.
[4] T. Levanen, J. Talvitie, and M. Renfors, "Performance Evaluation of a DDST Based SIMO SC System With PAPR Reduction," in *6th International Symposium on Turbo Codes & Iterative Information Processing, ISTC 2010,*, Brest, France, 6-10 Sept.
[5] S. He, J. K. Tugnait, and X. Meng, "On Superimposed Training for MIMO Channel Estimation and Symbol Detection," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 3007–3021, 2007.
[6] J. K. Tugnait and S. He, "Multiuser/MIMO Doubly Selective Fading Channel Estimation Using Superimposed Training and Slepian Sequences," *IEEE Trans. Veh. Technol.*, vol. 59, no. 3, pp. 1341–1354, 2010.
[7] T. Ihalainen et al., "Channel Equalization for Multi-Antenna FBMC/OQAM Receivers," *IEEE Trans. Veh. Technol.*, vol. 60, no. 5, pp. 2070–2085, June 2011.
[8] D. Chu, "Polyphase Codes with Good Periodic Correlation Properties," *IEEE Trans. Inf. Theory*, vol. 18, no. 4, pp. 531–532, Jul. 1072.
[9] A. G. Orozco-Lugo, M. M. Lara, and D. C. McLernon, "Channel Estimation Using Implicit Training," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 240–254, 2004.
[10] O. Weikert and U. Zölzer, "Efficient MIMO Channel Estimation With Optimal Tranining Sequences," in *in Proc. 1st Workshop on Commercial MIMO-Components and -Systems (CMCS 2007)*, Duisburg, Germany, 13-14 Sept. 2007.
[11] Y. Yang, T. Ihalainen, M. Rinne and M. Renfors, "Frequency-domain Equalization in Single-carrier Transmission: Filter Bank Approach," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, Article ID 10438, 16 pages, 2007.
[12] C. Laot, A. Glavieux, and J. Labat, "Turbo Equalization: Adaptive Equalization and Channel Decoding Jointly Optimized," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 9, pp. 1744–1752, 2001.
[13] T. Levanen, J. Talvitie and M. Renfors, "Improved Performance Analysis for Super Imposed Pilot Based Short Channel Estimator," in *IEEE International Workshop on Signal Processing Advances for Wireless Communications, SPAWC 2010*, Marrakech, Morocco, 20-23 June 2010.
[14] T. B. Sorensen, P. E. Mogensen, and F. Frederiksen, "Extension of the ITU Channel Models for Wideband (OFDM) Systems," in *62nd IEEE Vehicular Technology Conference, VTC 2005 Fall*, vol. 1, 25-28 Sept. 2005, pp. 392–396.
[15] C. Berrou, A. Glavieux and P. Thitimajshima, "Near Shannon Limit Error-Correcting Coding and Decoding: Turbo-Codes," in *IEEE International Conference on Communications*, vol. 2, Geneva, May 1993, pp. 1064–1070, conference Record.
[16] S. Sharma, S. Attri, and R.C. Chauhan, "A Simplified and Efficient Implementation of FPGA-Based Turbo Decoder," in *2003 IEEE International Performance, Computing, and Communications Conference*, 9-11 Apr. 2003, pp. 207–213.
[17] D. Divsalar and F. Pollara, "Turbo Codes for PCS Applications," in *IEEE International Conference on Communications, ICC'95*, Seattle, WA, USA, 18-22 June 1995, pp. 54–59.