Low-Complexity 2-Coordinates Descent for Near-Optimal MMSE Soft-Output Massive MIMO Uplink Data Detection

Pascal Seidel, Steffen Paul, Jochen Rust
Institute of Electrodynamics and Microelectronics (ITEM.me)
University of Bremen, Germany
{seidel, steffen.paul, rust}@me.uni-bremen.de

***Abstract*—In this paper, a block extended coordinate descent algorithm is introduced for MMSE based soft-output massive MIMO signal detection, which exploits the simple inversion of small sub-Gram matrices to allow a low-complexity implementation. We show that the resulting two-coordinates descent approach has a computational complexity comparable to the original coordinate descent signal detector, whereas the latency bottleneck can be relaxed and further the data detection performance can be improved as the simulation results show. Also we show the possibility to approximate the Gram matrix with fewer multiplications while maintaining a near-optimal detection performance.**

***Index Terms*—Massive MIMO, Soft-Output, Signal Detection, Low-Complexity, Matrix Approximation**

I. Introduction

High-performance wireless communication is an important issue of present research due to several ongoing trends, like cloud computing, industrial communications or vehicular communications (V2V) [1]. Beside the traditional performance objectives, like power consumption, complexity and throughput, robust communication has become a highly desirable constraint. For instance, in the scope of industrial radio, even small disturbances caused, e.g., by channel side effects or antenna correlations, must be minimized in order to avoid security-critical malfunction that may cause serious damages or human injuries.

A promising approach to achieve the ambitious objectives of future wireless communication systems is to go for concurrent multi-antenna communications, in particular massive multiple-input-multiple-output (MIMO) systems [2]. In this scope, very large antenna arrays are considered, where the number of antennas is scaled up to hundreds antennas in order to raise the robustness and/or the data rate. However, practically implementing a massive MIMO system quickly becomes a complicated task, due to several implementation challenges that have to be addressed like accurate antenna design, energy efficiency, connectivity and synchronization issues, etc. Above all, the efficient implementation of high-performance baseband signal detectors is of major interest.

To solve this problem, several algorithmic approaches have been investigated by now, mainly concentrating on iterative linear equalization methods based on estimation schemes like zero forcing (ZF) or minimum mean square error (MMSE), which achieve a near-optimal data detection when the basestation-to-client anntennas ratio (BCR) is large enough [2]. For instance, meaningful results can be achieved considering the iterative Richardson iterations [3], Conjugate-Gradient (CG) method [4] or the Gauss-Seidel (GS) approach [5]. However, these algorithms need the so-called Gram matrix and the matched filter output vector, whose determination incurs high computational complexity.

Another approach is based on the coordinate descent (CD) algorithm, where only the calculations of channel gains are required to determine the line-search step size; this allows a high-throughput soft-output signal estimation [6]. However, corresponding algorithms suffer from a high latency because it updates the estimated signals for each user in a strict, sequential manner.

In this paper, we propose an extension to the MMSE-based CD soft-output detector by performing an iterative signal estimation for two transmit signals at once, where a simple inversion of 2×2 sub-Gram matrices is applied. The two-coordinates descent (2CD) outperforms the traditional approach but has a comparable low complexity. An approximation methodology for Gram matrices is proposed to reduce the number of computations for the preprocessing stage, which is also applicable to existing iterative signal detection methods.

II. Preliminaries

We consider a centralized large-scale MIMO uplink system where a number of N_{C} non-cooperative single-antenna clients simultaneously transmit digitally modulated data over a wireless channel to a basestation (BS) consisting of $N_{\mathrm{B}} \gg N_{\mathrm{C}}$ receiving antennas. Furthermore, the wireless channel is assumed to be flat, e.g. by using orthogonal frequency division multiplexing (OFDM) with a sufficient cyclic prefix, so the received signal at the BS $\mathbf{y} \in \mathbb{C}^{N_{\mathrm{B}}}$ can be described by the linear equation system

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}\,, \tag{1}$$

where the vector $\mathbf{s} \in \mathcal{A}^{N_{\mathrm{C}}}$ represents the compilation of all client-specific transmitted symbols s_i with $i = 1, \ldots, N_{\mathrm{C}}$, taken from a fixed normalized modulation alphabet $\mathcal{A}$. Under the assumption of a flat Rayleigh fading, the entries of the wireless channel matrix $\mathbf{H} \in \mathbb{C}^{N_{\mathrm{B}} \times N_{\mathrm{C}}}$ are symmetric

978-9-0827-9703-9/19/$31.00 ©2019 IEEE

complex Gaussian, independently and identically distributed (i.i.d) with zero mean and unit variance. The additive white Gaussian noise vector $\mathbf{n} \in \mathbb{C}^{N_\mathrm{B}}$ holds i.i.d. entries with zero mean and variance σ_n^2. Further, perfectly known channel state information (CSI) is assumed. Considering the channel model from eq. (1), the optimal symbol detection can be achieved by solving the maximum-likelihood (ML) problem

$$\hat{\mathbf{s}}_\mathrm{ML} = \underset{\mathbf{s}\in\mathcal{A}^{N_\mathrm{C}}}{\arg\min} \|\mathbf{Hs} - \mathbf{y}\|_2^2 \, , \tag{2}$$

which requires an exhaustive search through all possible combinations of symbols in $\mathcal{A}$. Due to the large number of linear equations in eq. (1), the ML estimation approach is impractical for massive MIMO systems, as computational complexity grows exponentially. Also, even approximate ML methods like sphere detection or lattice reduction could be implemented only for highly limited modulation alphabets [7] [8]. On the other hand, by allowing solutions all over the complex vector space $\hat{\mathbf{s}} \in \mathbb{C}^{N_\mathrm{C}}$, linear detector schemes like ZF or MMSE can be applied, which have a much lower complexity compared to the ML method. Under the assumption that the BS has knowledge of the noise power σ_n^2 the MMSE detection is given by the solution of the following unconstrained optimization problem:

$$\hat{\mathbf{s}}_\mathrm{MMSE} = \underset{\mathbf{s}\in\mathbb{C}^{N_\mathrm{C}}}{\arg\min} \|\mathbf{Hs} - \mathbf{y}\|_2^2 + \sigma_\mathrm{n}^2 \|\mathbf{s}\|_2^2 \tag{3}$$

Equivalently, the estimated signal vector $\hat{\mathbf{s}}_\mathrm{MMSE}$ from eq. (3) can also be computed directly by the closed-form solution of the regularized least squares problem

$$\hat{\mathbf{s}}_\mathrm{MMSE} = \left(\mathbf{H}^H\mathbf{H} + \sigma_\mathrm{n}^2\mathbf{I}_{N_\mathrm{C}}\right)^{-1}\mathbf{H}^H\mathbf{y} = \tilde{\mathbf{G}}^{-1}\mathbf{y}_\mathrm{MF} \, , \tag{4}$$

where it is required to determine the inverse of the regularized Hermitian Gram matrix $\tilde{\mathbf{G}} = \mathbf{H}^H\mathbf{H} + \sigma_\mathrm{n}^2\mathbf{I}_{N_\mathrm{C}}$ and the matched filter output vector $\mathbf{y}_\mathrm{MF} = \mathbf{H}^H\mathbf{y}$. When we can assume that the estimate in eq. (4) is close to the ML solution, reliability information like log-likelihood ratios (LLR) for a soft-input channel decoder can be extracted, which results in a huge data detection performance gain [2]. The max-log approximated LLR L_b of bit b for the i-th client signal estimate $\hat{s}_i$ is then given by:

$$L_b(\hat{s}_i) = \rho_i \left(\min_{c\in\mathcal{A}_b^0} \left| \frac{\hat{s}_i}{\mu_i} - c \right|^2 - \min_{c\in\mathcal{A}_b^1} \left| \frac{\hat{s}_i}{\mu_i} - c \right|^2 \right) \tag{5}$$

The subsets $\mathcal{A}_b^0$ and $\mathcal{A}_b^0$ contain the constellation points of the alphabet $\mathcal{A}$ where the bit b is equal to 0 and 1, respectively. When we define the equivalent post-equalization channel matrix as $\mathbf{E} = \tilde{\mathbf{G}}^{-1}\mathbf{G}$ with the Gram matrix $\mathbf{G} = \mathbf{H}^H\mathbf{H}$, the necessary post-equalization channel gain and signal-to-interference-plus-noise ratio (SINR) are given by $\mu_i = e_{ii}$ and $\rho_i = \mu_i/(1-\mu_i)$, respectively. For large BCR $\varrho = N_\mathrm{B}/N_\mathrm{C}$ it was shown that the MMSE detection method achieves a near-ML data detection performance regarding to the bit error rate [2]. Nevertheless, the computational complexity of computing the Gram matrix $\mathbf{G}$ and its regularized inverse $\tilde{\mathbf{G}}^{-1}$ for the closed-form solution of the MMSE estimation problem in eq. (4) is $\mathcal{O}(N_\mathrm{B}N_\mathrm{C}^2)$ and $\mathcal{O}(N_\mathrm{C}^3)$, which becomes an infeasible task for a continuously growing number of antennas in a massive MIMO communication system. To avoid the explicit inversion of $\tilde{\mathbf{G}}$, various approaches like the inverse Gram matrix approximation by a Neumann-series or iterative algorithms like Richardson iterations, Gauss-Seidel method or the conjugate gradient algorithm are able to reduce the complexity to $\mathcal{O}(N_\mathrm{C}^2)$ but still require to compute the Gram matrix $\mathbf{G}$, which is hard to realize for scenarios where the wireless channel is assumed to change rapidly [3] [4] [5] [9].

III. Two-Coordinates Descent based MMSE Soft-Output data detection

A promising approach for the soft-output signal detection avoiding the explicit computation of the full Gram matrix is the CD method that restricts the minimization of the target function in eq. (3) along the direction of a single component of $\mathbf{s}$ in a cyclic iterative fashion. While this allows a high-throughput hardware realization, the latency is comparatively high compared to normal equations based on approaches because of the strongly sequential implementation [6]. To tackle this problem we propose an extension to the classical CD approach by introducing coordinate blocks instead of single coordinates. By choosing these blocks small enough and using further subsampling approximations to reduce the number of scalar product operations, the complexity of the optimization algorithm does not increase significantly while the sequential character of the classical CD algorithm is relaxed.

A. Generalized block coordinate descent

The derivation of the block extended version of the CD algorithm is similar to the original approach in [6]. First, we set the gradient of the cost function of eq. (3), denoted as $\varphi(\mathbf{s})$, in respect to a subset of components $\mathbf{s}_\mathcal{B}$ with $\mathcal{B} \subset \{1, \ldots, N_\mathrm{C}\}$ to zero to find the optimal vector $\hat{\mathbf{s}}_\mathcal{B}$.

$$\mathrm{grad}_\mathcal{B}(\varphi(\mathbf{s})) = \mathbf{H}_\mathcal{B}^H(\mathbf{Hs} - \mathbf{y}) + \sigma_\mathrm{n}^2\mathbf{s}_\mathcal{B} = 0 \tag{6}$$

By splitting the channel matrix $\mathbf{H} = [\mathbf{H}_\mathcal{B}|\mathbf{H}_\mathcal{I}]$ and the symbol vector $\mathbf{s}^H = \left[\mathbf{s}_\mathcal{B}^H|\mathbf{s}_\mathcal{I}^H\right]$ with $\mathcal{I} = \{1, \ldots, N_\mathrm{C}\} \setminus \mathcal{B}$, the solution of eq. (6) can be written as

$$\hat{\mathbf{s}}_\mathcal{B} = \tilde{\mathbf{G}}_\mathcal{B}^{-1}\mathbf{H}_\mathcal{B}^H(\mathbf{y} - \mathbf{H}_\mathcal{I}\mathbf{s}_\mathcal{I}) \, , \tag{7}$$

where $\tilde{\mathbf{G}}_\mathcal{B} = \mathbf{H}_\mathcal{B}^H\mathbf{H}_\mathcal{B} + \sigma_\mathrm{n}^2\mathbf{I}$ is the so-called Hermitian regularized sub-Gram matrix, with $\mathbf{H}_\mathcal{B} \in \mathbb{C}^{N_\mathrm{B}\times|\mathcal{B}|}$ containing the columns of $\mathbf{H}$ regarding the elements in subset $\mathcal{B}$. Like in the original formulation of the MMSE based CD soft-output detector, we define a residual approximation vector $\mathbf{r}^{(n)} \in \mathbb{C}^{N_\mathrm{B}}$ for the n-th iteration:

$$\mathbf{r}^{(n)} = \mathbf{y} - \mathbf{H}_{\mathcal{I}_n}\hat{\mathbf{s}}_{\mathcal{I}_n}^{(n)} - \mathbf{H}_{\mathcal{B}_n}\hat{\mathbf{s}}_{\mathcal{B}_n}^{(n)} \tag{8}$$

so that eq. (7) can be rewritten for the n-th iteration

$$\hat{\mathbf{s}}_{\mathcal{B}_n}^{(n)} = \tilde{\mathbf{G}}_{\mathcal{B}_n}^{-1}\left(\mathbf{H}_{\mathcal{B}_n}^H\mathbf{r}^{(n-1)} + \mathbf{G}_{\mathcal{B}_n}\hat{\mathbf{s}}_{\mathcal{B}_n}^{(n-1)}\right) \tag{9}$$

with the initialization $\hat{\mathbf{s}}^{(0)} = \mathbf{0}$ and $\mathbf{r}^{(0)} = \mathbf{y}$. Note, that the component subsets $\mathcal{B}_n$ and $\mathcal{I}_n$ vary for each iteration to ensure that the minimization of $\varphi(\mathbf{s})$ is performed at least once for each component. Due to the use of the sub-Gram matrix in eq. (9) the block coordinate descent (BCD) algorithm also keeps the non-normalized inter-client interference coefficients $g_{\mathcal{B}_n,ij}$ for the soft-output detection into account, which results in a superior convergence rate compared to the classical CD algorithm.

B. Two-coordinates descent

In comparison to the classical CD approach, the block extended method requires the inversion of the sub-Gram matrices, which incurs significant computational complexity when the coordinate blocks are large. But fortunately, when the blocksize is restricted to $|\mathcal{B}| = 2$, the inverse of the sub-Gram matrix can be calculated explicitly at low complexity by $\tilde{\mathbf{G}}_{\mathcal{B}}^{-1} = \gamma_{\mathcal{B}}^{-1}\tilde{\mathbf{C}}_{\mathcal{B}}$ with $\tilde{\mathbf{C}}_{\mathcal{B}} = \mathrm{adj}(\tilde{\mathbf{G}}_{\mathcal{B}})$ and $\gamma_{\mathcal{B}} = \det(\tilde{\mathbf{G}}_{\mathcal{B}})$ being the adjugate and determinant of the regularized sub-Gram matrix $\tilde{\mathbf{G}}_{\mathcal{B}}$, respectively. In that specific case $\tilde{\mathbf{C}}_{\mathcal{B}}$ and $\gamma_{\mathcal{B}}^{-1}$ are given by:

$$\tilde{\mathbf{C}}_{\mathcal{B}} = \begin{pmatrix} g_{\mathcal{B},22} + \sigma_{\mathrm{n}}^2 & -g_{\mathcal{B},12} \\ -g_{\mathcal{B},12}^* & g_{\mathcal{B},11} + \sigma_{\mathrm{n}}^2 \end{pmatrix} = \mathbf{C}_{\mathcal{B}} + \sigma_{\mathrm{n}}^2\mathbf{I} \tag{10}$$

$$\gamma_{\mathcal{B}}^{-1} = \left(g_{\mathcal{B},11}g_{\mathcal{B},22} - |g_{\mathcal{B},12}|^2 + \sigma_{\mathrm{n}}^2(g_{\mathcal{B},11} + g_{\mathcal{B},22} + \sigma_{\mathrm{n}}^2)\right)^{-1} \tag{11}$$

As it can be seen in eq. (10) due to the addition of the noise variance σ_{n}^2 on the main diagonal of the sub-Gram matrix $\mathbf{G}_{\mathcal{B}}$, the adjugate of $\tilde{\mathbf{G}}_{\mathcal{B}}$ can be alternatively described by a sum of the adjugate of the sub-Gram matrix $\mathbf{C}_{\mathrm{B}}$ and the diagonal MMSE regularization matrix $\sigma_{\mathrm{n}}^2\mathbf{I}$. In that specific case $\tilde{\mathbf{C}}_{\mathcal{B}}$ can be determined without further computations, as is it given by simple element interchanging of $\mathbf{G}_{\mathcal{B}}$ and sign flipping. Additionally, the calculations of the scaling factors $\gamma_{\mathcal{B}}^{-1}$ in eq. (11) contain only real valued multiplications because of the Hermitian structure of the regularized sub-Gram matrix. Using eq. (10) and (11) the inverse of the regularized sub-Gram matrix $\tilde{\mathbf{G}}_{\mathcal{B}}$ can be explicitly computed by:

$$\tilde{\mathbf{G}}_{\mathcal{B}}^{-1} = \gamma_{\mathcal{B}}^{-1}\left(\mathbf{C}_{\mathcal{B}} + \sigma_{\mathrm{n}}^2\mathbf{I}\right) \tag{12}$$

Using the explicit inverse in eq. (12) for the BCD update equation with $|\mathcal{B}|$ in eq. (9) leads to the update equations of the two-coordinates descent (2CD) algorithm:

$$\hat{\mathbf{z}}_t^{(k)} = \mathbf{H}_{\mathcal{B}_t}^H\mathbf{r}_{t-1}^{(k)} + \mathbf{G}_{\mathcal{B}_t}\hat{\mathbf{s}}_{\mathcal{B}_t}^{(k-1)} \tag{13}$$

$$\hat{\mathbf{s}}_{\mathcal{B}_t}^{(k)} = \gamma_{\mathcal{B}_t}^{-1}\left[\mathbf{C}_{\mathcal{B}_t}\hat{\mathbf{z}}_t^{(k)} + \sigma_{\mathrm{n}}^2\hat{\mathbf{z}}_t^{(k)}\right] \tag{14}$$

$$\mathbf{r}_t^{(k)} = \mathbf{r}_{t-1}^{(k)} - \mathbf{H}_{\mathcal{B}_t}\left(\hat{\mathbf{s}}_{\mathcal{B}_t}^{(k)} - \hat{\mathbf{s}}_{\mathcal{B}_t}^{(k-1)}\right) \tag{15}$$

To reduce the number of computations, a cyclic selection scheme $\mathcal{B}_t = \{2t-1, 2t\}$ for $t = 1, \ldots, N_{\mathrm{C}}/2$ is used, following that each iteration k consists of updating each estimate $\hat{s}_i$ once. Due to the fixed component sets $\mathcal{B}_t$, the sub-Gram matrices can be precomputed and reused in each iteration k. After a number of K iterations, the estimates $\hat{s}_i$ are used to calculate the LLR with the approximation from [4] for the post-equalization channel gains $\tilde{\mu}_i = \|\mathbf{h}_i\|_2^2/\left(\|\mathbf{h}_i\|_2^2 + \sigma_n^2\right)$ and the post-equalization SINR $\tilde{\rho}_i = \tilde{\mu}_i/(1-\tilde{\mu}_i)$, where $\|\mathbf{h}_i\|_2^2$ denotes the squared norm of the i-th column of the channel matrix $\mathbf{H}$.

TABLE I: Number of real valued multiplications

	Initialization	K time iteration
RI [3]	$2N_{\mathrm{B}}N_{\mathrm{C}}^2$	$(4N_{\mathrm{C}}^2 + 2N_{\mathrm{C}})K$
CG [4]	$2N_{\mathrm{B}}N_{\mathrm{C}}^2$	$(4N_{\mathrm{C}}^2 + 10N_{\mathrm{C}})K$
GS [5]	$2N_{\mathrm{B}}N_{\mathrm{C}}^2$	$4N_{\mathrm{C}}^2K$
CD [6]	$2N_{\mathrm{B}}N_{\mathrm{C}} + N_{\mathrm{C}}$	$(8N_{\mathrm{B}}N_{\mathrm{C}} + 4N_{\mathrm{C}})K$
2CD	$4N_{\mathrm{B}}N_{\mathrm{C}} + 3N_{\mathrm{C}}$	$(8N_{\mathrm{B}}N_{\mathrm{C}} + 14N_{\mathrm{C}})K$

C. Sub-Gram matrix approximation

One of the main advantages of the coordinate descend algorithms is the relatively low overhead, i.e. these methods do not require the computation of the full Gram matrix $\mathbf{G}$ and the matched filter vector $\mathbf{y}_{\mathrm{MF}}$ in eq. (4). Instead, for the CD and 2CD we just have to determine the channel gains and the sub-Gram matrix with their associated inverse determinant as a preprocessing step [6]. To further reduce the number of computations in the preprocessing stage of the 2CD we can approximate the sub-Gram matrices, which are used in the update equations (13) and (14). As it was shown for a flat fading Rayleigh channel model the channel matrix $\mathbf{H}$ can be approximated regarding their singular values by a simple randomized row subsampling [10]. By adopting this idea, the sub-Gram matrices for the proposed 2CD can be approximated by

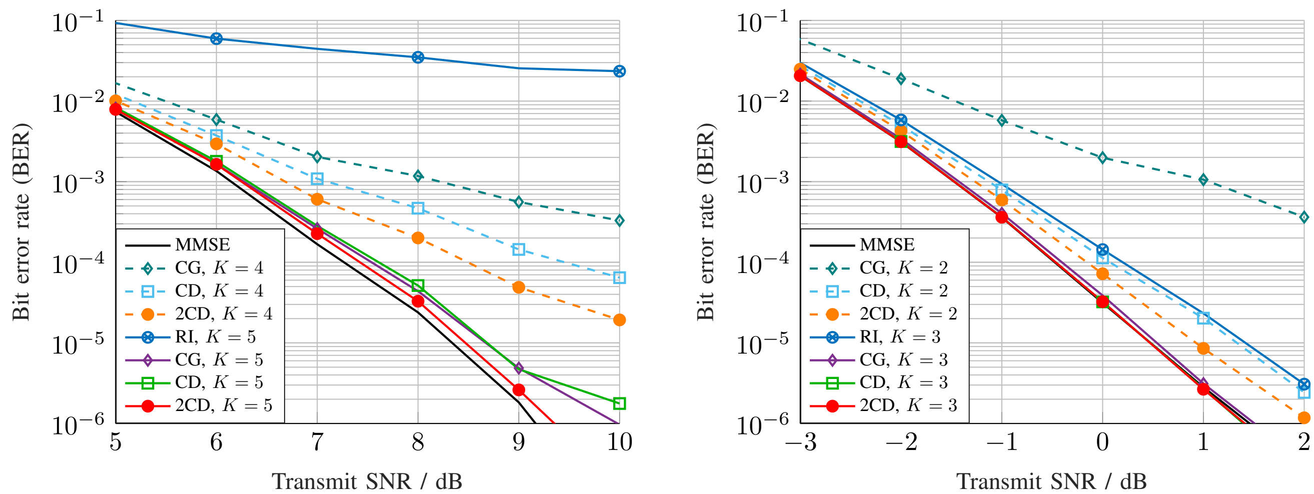
$$\mathbf{G}_{\mathcal{B}_t} \approx \frac{1}{\eta}\left(\mathbf{S}\mathbf{H}_{\mathcal{B}_t}\right)^H\left(\mathbf{S}\mathbf{H}_{\mathcal{B}_t}\right), \tag{16}$$

with $0 < \eta \leq 1$ the subsampling factor and the so-called subsampling matrix $\mathbf{S} \in \{0,1\}^{\eta N_{\mathrm{B}} \times N_{\mathrm{B}}}$, which holds a random subset of ηN_{B} rows from an $N_{\mathrm{B}} \times N_{\mathrm{B}}$ identity matrix. If we assume that each antenna at the BS receives approximatively the same energy, the scaling factor $1/\eta$ compensates the power loss caused by the subsampling procedure. Note that for the CD algorithm, the same method can be applied to approximate the channel gains, whose reciprocal values are used as constant step sizes.

IV. Evaluation

A. Computational complexity analysis

For a comparison of the complexity between the proposed 2CD algorithm and the normal-equations based methods Richardson iterations (RI) [3], Gauss-Seidel (GS) method [5], conjugate gradient (CG) method [4] and the original coordinate descent (CD) approach [6], we count the number of real valued multiplications for a fixed number of iterations K in dependency of the number of BS antennas N_{B} and the single-antenna clients N_{C}, which are shown in table I. We assume that one complex valued multiplication corresponds to four real valued multiplications. We differentiate between initialization complexity, where e.g. the sub-Gram matrices are calculated, and iteration dependent complexity, which scales with the number of total iterations K for the successive detection procedure. Various computations for initial estimates and the LLR calculates are not included. In the comparison,



(a) 32 BS antennas $\times$ 8 clients

(b) 128 BS antennas $\times$ 8 clients

Fig. 1: Bit error rate (BER) for a massive MIMO system with $N_\mathrm{C} = 8$ single-antenna clients using the iterative soft-output detectors RI [3], CG [4], CD [6] and our proposed 2CD at different transmit signal-to-noise ratios (SNR) with a varying number of iterations K for (a) $N_\mathrm{B} = 32$ and (b) $N_\mathrm{B} = 128$ antennas at the BS. The optimal MMSE detector (black line) is given as reference.

the initialization complexity for the CD and 2CD approaches scales linear to the number of antennas at the BS N_B and the number of single-antenna clients N_C whereas the number of real valued multiplications increase linear to N_B and quadratically with N_C for the normal-equation based detectors. This evolves to a critical state when the number of clients in the communication system grows.

Comparing the CD with our proposed 2CD it can been seen that our approach requires more real valued multiplications for the initialization, but has roughly the same complexity in the iterative processing stage. Note that due to the parallelization of the CD method in form of the 2CD the slightly increased computational complexity in the preprocessing is unavoidable.

B. BER performance

To evaluate the data detection performance of the proposed 2CD based soft-output MMSE detector we evaluate the bit error rate (BER) performance for different transmit signal-to-noise ratios (SNR). For the simulations we use the channel model from eq. (1) with $N_\mathrm{C} = 8$ and $N_\mathrm{B} = \{32, 128\}$ two different BCR ϱ. In the simulations, each client uses a punctured rate 5/6 convolutional code with polynomials $[133_o, 171_o]$ and 64-QAM modulation. At the receiver side, the extracted LLR from the soft-output detector is decoded using a soft-input Viterbi decoder.

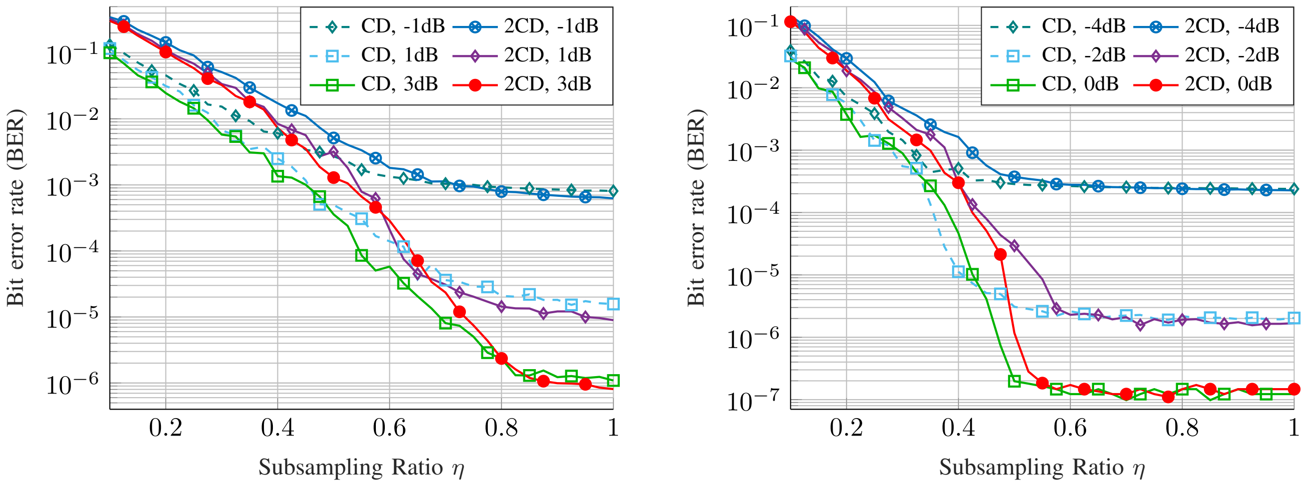
The BER for the proposed soft-output detector is shown in fig. 1 for both system dimensions with a fixed number of iterations K. For a comparison to existing methods, we evaluate also the approximate MMSE detectors RI, CG and CD, where no further initialization is used. Additionally, we use the approximation for the post-equalization channel gains and SINR from section III-B. Note that we exclude the GS method from the simulations, as it is just a special implementation of the cyclic CD based on the normal equations [11].

In case of $N_\mathrm{B} = 32$ antennas at the BS ($\varrho = 4$) the CG, CD and 2CD approaches achieve a close-to MMSE performance for $K = 4$ iterations and a near-optimal performance for $K = 5$ iterations. As it is shown in fig. 1a the 2CD approach significantly outperforms the CG and CD method for $K = 4$. By increasing the number of iterations to $K = 5$ the performance gap between the iterative method CG, CD method and the optimal MMSE detector gets larger for higher SNR, while our 2CD approach still achieves a near-MMSE BER performance. Due to the use of the inter-client interference information in the 2CD algorithm, the proposed method can still attain a near-MMSE BER performance within a limited number of iterations, even when the BCR ϱ becomes small.

When the BCR increases to $\varrho = 16$ with $N_\mathrm{B} = 128$ the necessary number of iterations for the low-complexity soft-output detection methods reduces, which can be seen in fig. 1b. While the 2CD method noticeably outperforms CG and CD at $K = 2$ iterations, all of them achieve a near-MMSE BER performance for $K = 3$. As the client-channels in the massive MIMO become nearly uncorrelated for high ϱ due to the channel hardening property, the absolute values of the off-diagonal elements of the sub-Gram matrix decrease as well, so that the convergence rate of the 2CD becomes similar to the CD method.

C. Subsampling capabilities

As shown in sec. IV-A the initialization of the CD and 2CD method has a significantly lower computational complexity compared to the normal equation based approaches as these algorithms do not require the full Gram matrix and matched filter output vector. Especially for high BCR ϱ, where the



(a) 128 BS antennas × 8 clients

(b) 256 BS antennas × 8 clients

Fig. 2: Bit error rate (BER) for a massive MIMO system with $N_\mathrm{C} = 8$ single-antenna clients using the soft-output detectors CD [6] and our proposed 2CD with different subsampling ratios η for (a) $N_\mathrm{B} = 128$ and (b) $N_\mathrm{B} = 256$ antennas at the BS. The BER is evaluated at three different transmit signal-to-noise ratios (SNR).

linear detection methods achieve a near-MMSE at minimum complexity, the latency, caused by the preprocessing, becomes critical. To mitigate this bottleneck, we use the subsampling methodology in sec. III-C to approximate the channel gains and sub-Gram matrices with fewer multiplications for the CD and 2CD detectors, respectively. To investigate the impact of the subsampling based approximation we evaluate the BER with the simulation parameters from sec. IV-B for $N_\mathrm{B} = \{128, 256\}$ and $K = 2$ over the subsampling parameter η at three different SNR, which is shown in fig. 2.

For $N_\mathrm{B} = 128$ ($\varrho = 16$) it can be seen in fig. 2a that the BER of both detection methods decrease continuously by increasing the subsampling parameter η. Interestingly, for $\eta \geq 0.9$ the algorithms achieve a BER close to the non-approximation case ($\eta = 1$). While the approximation of the channel gains in the CD method has a weaker degradation for $\eta < 0.9$, the 2CD outperforms this method in the stable region mentioned before, still.

The same BER characteristic can be seen in fig. 2b, where the number of BS antennas is increased to $N_\mathrm{B} = 256$ ($\varrho = 32$). In comparison the BER decreases rapidly when the subsampling factor increases, where for $\eta > 0.6$ the optimal signal detection performance for both methods can be accomplished.

V. Conclusion

In this paper, we extended the original CD soft-output detector to the block CD, which performs the MMSE optimization over multiple transmit signals at once. By restricting the block size to two, a simple inversion of the resulting sub-Gram matrices could be achieved, with a computational complexity close to the CD detector, whereas the data detection performance outperforms the original approach. Furthermore, it was shown that for very high BCR the initialization of the CD and 2CD methods can be simplified by random subsampling without noticeable performance loss in the signal detection.

References

[1] Jeffrey G Andrews, Stefano Buzzi, Wan Choi, Stephen V Hanly, Angel Lozano, Anthony CK Soong, and Jianzhong Charlie Zhang, "What will 5G be?," *IEEE Journal on selected areas in communications*, vol. 32, no. 6, pp. 1065–1082, 2014.

[2] Fredrik Rusek et al., "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, 2013.

[3] Xinyu Gao, Linglong Dai, Yongkui Ma, and Zhaocheng Wang, "Low-complexity near-optimal signal detection for uplink large-scale MIMO systems," *Electronics letters*, vol. 50, no. 18, pp. 1326–1328, 2014.

[4] Bei Yin, Michael Wu, Joseph R Cavallaro, and Christoph Studer, "Conjugate gradient-based soft-output detection and precoding in massive MIMO systems," in *Global communications conference (GLOBECOM), 2014 IEEE*. IEEE, 2014, pp. 3696–3701.

[5] Linglong Dai et al., "Low-complexity soft-output signal detection based on Gauss-Seidel method for uplink multiuser large-scale MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 10, 2015.

[6] Michael Wu, Chris Dick, Joseph R Cavallaro, and Christoph Studer, "High-throughput data detection for massive MU-MIMO-OFDM using coordinate descent," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 63, no. 12, pp. 2357–2367, 2016.

[7] Yo-Seb Jeon, Namyoon Lee, Song-Nam Hong, and Robert W Heath, "One-bit sphere decoding for uplink massive MIMO systems with one-bit ADCs," *IEEE Transactions on Wireless Communications*, 2018.

[8] Qi Zhou and Xiaoli Ma, "Element-based lattice reduction algorithms for large MIMO detection," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 274–286, 2013.

[9] Bei Yin, Michael Wu, Christoph Studer, Joseph R Cavallaro, and Chris Dick, "Implementation trade-offs for linear detection in large-scale MIMO systems," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE, 2013, pp. 2679–2683.

[10] Pascal Seidel, Benjamin Knoop, Sebastian Schmale, Daniel Gregorek, Steffen Paul, and Jochen Rust, "Random subsampling based signal detection for spatial correlated massive MIMO channels," in *Circuits and Systems (ISCAS), 2018 IEEE International Symposium on*. IEEE, 2018, pp. 1–5.

[11] Amir Beck and Luba Tetruashvili, "On the convergence of block coordinate descent type methods," *SIAM journal on Optimization*, vol. 23, no. 4, pp. 2037–2060, 2013.