ALGORITHMIC SOLUTIONS FOR PILOT DESIGN OPTIMIZATION IN ARBITRARILY CORRELATED SCENARIOS

Mats Bengtsson

Signal Processing Department, ACCESS Linnaeus Center, KTH Royal Institute of Technology, Sweden

ABSTRACT

We consider optimized design of training sequences, given knowledge of the channel and noise statistics. Recently, pilot designs considering the end performance of the channel estimate, have been proposed, both optimizing the average performance and the performance at a certain outage level. Unfortunately, these problems, as well as previously proposed designs optimizing the channel estimation MSE, are non-convex for arbitrary channel and noise correlations so additional assumptions have been introduced in the literature to be able to find tractable solutions. Here, we show that arbitrarily correlated scenarios can easily be handled by resorting to alternating optimization, for all the previously mentioned problem formulations. Furthermore, we numerically compare the average and outage performance of the proposed algorithms, to alternative solutions adopted from the literature.

1. INTRODUCTION

Channel state information (CSI) in some form is needed in almost all wireless communication systems, especially in MIMO systems. Primarily, the CSI is used at the receiver to equalize and detect the incoming data, but it can also be used at the transmitter side for precoding and scheduling, for example. MIMO channel estimation is typically done using pilot sequences that are spatially white, which has been shown to be the best choice according to several criteria [1, 2]. However, if prior knowledge is available about the channel state information, the pilot signal can be adapted to improve the estimation accuracy. Such optimized designs have among others been considered for correlated channels in white noise [3, 4], for uncorrelated channels in colored noise [5] and for the general case of correlated channels in colored noise [6, 7, 8]. All work mentioned so far considers the mean square error (MSE) of the resulting channel estimate as the optimization criterion. A few other criteria have also been proposed, such as optimizing the mutual information between the CSI and the received training data [9].

Minimizing the MSE of the channel estimate does not necessarily provide the optimal end user performance. Inspired by the system identification literature, such as [10], the recent paper [11] shows how to incorporate the application in the pilot design, using a cost function that reflects the impact of the channel estimation error on the end performance. Two main strategies are proposed, one related to the average performance and one related to the probability that the performance is good enough.

A common problem of all the mentioned pilot design strategies is that an easily computable solution only is available under certain conditions on Kronecker structures both in the channel and noise covariance matrices, that may be questionable in real-world scenarios. Here, we show that the optimization problems resulting from the different pilot design strategies can be solved for general (also non-Kronecker) correlation structures, using alternating optimization. To illustrate the pros and cons of this alternating optimization approach, we provide numerical comparisons to a number of alternatives from the literature, both to optimal solutions for scenarios where such can be found and to reasonable engineering solutions for scenarios where no optimal solution can be found in the literature.

This work was supported by the FP7 project HIATUS (FET-Open grant #265578).

2. PRELIMINARIES

We consider a single frequency flat MIMO link (e.g. one subcarrier in an OFDM system), with N_t transmit and N_r receive antennas,

$$\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{n}(t) \ , \tag{1}$$

with transmitted signal $\mathbf{x}(t)$, received signal $\mathbf{y}(t)$ and additive interference plus noise $\mathbf{n}(t)$. During the training phase, a length B pilot signal, $\mathbf{p}(1), \mathbf{p}(2), \ldots, \mathbf{p}(B)$, is transmitted. Collecting the B transmitted pilot vectors into a matrix $\mathbf{P} \in \mathbb{C}^{N_t \times B}$, the total received signal during the training phase is

$$\mathbf{Y} = \mathbf{H}\mathbf{P} + \mathbf{N} \ . \tag{2}$$

Assuming that the long-term channel statistics is known as $\mathbf{h} = \mathrm{vec}\{\mathbf{H}\} \in \mathrm{CN}(0, \mathbf{R})$ and similarly that the interference plus noise statistics is $\mathrm{vec}\{\mathbf{N}(t)\} \in \mathrm{CN}(0, \mathbf{S})$, we form a linear channel estimate of the form

$$\hat{\mathbf{h}} = \mathrm{vec}\{\hat{\mathbf{H}}\} = \mathbf{W}\,\mathrm{vec}\{\mathbf{Y}\} = \mathbf{W}(\tilde{\mathbf{P}}\mathbf{h} + \mathbf{n}) \ , \tag{3}$$

where $\tilde{\mathbf{P}} = \mathbf{P}^T \otimes \mathbf{I}$, which results in an error covariance matrix of

$$\begin{aligned} \mathbf{C}_{\text{MSE}} &= \mathrm{E}[(\hat{\mathbf{h}} - \mathbf{h})(\hat{\mathbf{h}} - \mathbf{h})^H] \\ &= (\mathbf{W}\tilde{\mathbf{P}} - \mathbf{I})\mathbf{R}(\mathbf{W}\tilde{\mathbf{P}} - \mathbf{I})^H + \mathbf{W}\mathbf{S}\mathbf{W}^H \end{aligned} \tag{4}$$

In particular, the MMSE channel estimate

$$\mathbf{W} = (\mathbf{R}^{-1} + \tilde{\mathbf{P}}^H\mathbf{S}^{-1}\tilde{\mathbf{P}})^{-1}\tilde{\mathbf{P}}^H\mathbf{S}^{-1} \ , \tag{5}$$

gives the MMSE covariance matrix

$$\mathbf{C}_{\text{MSE}} = (\mathbf{R}^{-1} + \tilde{\mathbf{P}}^H\mathbf{S}^{-1}\tilde{\mathbf{P}})^{-1} \ . \tag{6}$$

2.1. Problem Formulations

Our goal is to determine the optimal training sequence $\mathbf{P}$, given a constraint on the total training energy. However, depending on how the channel estimate will be used, it is not necessarily optimal to aim for a minimal channel estimate MMSE. Instead, we follow [11] and consider an end performance metric $M(\hat{\mathbf{H}})$ of the equalizer or

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

precoder or whatever application the channel estimate is intended for and define our cost metric as the difference in end performance between using the channel estimate and using the true channel,

$$J(\tilde{\mathbf{h}}) = M(\hat{\mathbf{H}}) - M(\mathbf{H})\ . \tag{7}$$

As is shown in [11], this cost metric can often be well approximated by a quadratic expression of the form

$$J(\tilde{\mathbf{h}}) = \tilde{\mathbf{h}}^H \boldsymbol{\mathcal{I}}_{\text{adm}} \tilde{\mathbf{h}}\ . \tag{8}$$

Since $J(\tilde{\mathbf{h}})$ itself is stochastic, we study the following optimization criteria.

Average Performance Optimize the average performance

$$\begin{aligned} \min_{\mathbf{P}} \quad & \mathrm{E}[J(\tilde{\mathbf{h}})] = \mathrm{Tr}[\boldsymbol{\mathcal{I}}_{\text{adm}} \mathbf{C}_{\text{MSE}}] \\ \text{s.t.} \quad & \|\mathbf{P}\|_F^2 \le P_{\max} \end{aligned} \tag{9}$$

Outage Performance Maximize the chance α that the estimation error is inside an **admissible region** $J(\tilde{\mathbf{h}}) \le 1/\gamma$, for some desired accuracy γ.

$$\begin{aligned} \max_{\mathbf{P},\alpha} \quad & \alpha \\ \text{s.t.} \quad & \Pr[J(\tilde{\mathbf{h}}) \le 1/\gamma] = \alpha \\ & \|\mathbf{P}\|_F^2 \le P_{\max} \end{aligned} \tag{10}$$

Confidence Ellipsoids Since the outage formulation is hard to handle, we resort to bounding techniques. An approach proposed in [10, 11] is to define an ellipsoidal confidence region for which the probability is easy to calculate, $\mathcal{D} = \{\tilde{\mathbf{h}}^H \mathbf{C}_{\text{MSE}}^{-1} \tilde{\mathbf{h}} \le \frac{1}{2}\chi^2_\alpha(2N_tN_r)\}$[1] and replace the constraint $\Pr[J(\tilde{\mathbf{h}}) \le 1/\gamma] = \alpha$ by the tighter constraint that $J(\tilde{\mathbf{h}}) \le 1/\gamma$ for all $\tilde{\mathbf{h}} \in \mathcal{D}$. As shown in [11], the resulting pilot design problem can be written as

$$\begin{aligned} \max_{\mathbf{P},c} \quad & c \\ \text{s.t.} \quad & \mathbf{C}_{\text{MSE}}^{-1} \succeq c\boldsymbol{\mathcal{I}}_{\text{adm}} \\ & \|\mathbf{P}\|_F^2 \le P_{\max}\ , \end{aligned} \tag{11}$$

where c is related to $alpha$ through $c = \gamma\chi^2_\alpha(2N_tN_r)/2$. An alternative that often works better for practically useful outage levels α is to use a Markov bound to turn (10) into (9), see [12] for details.

2.2. Solutions from the Literature

Most existing literature considers the channel estimate MSE $\mathrm{Tr}[\mathbf{C}_{\text{MSE}}]$ as the optimization criterion. Even for this special case ($\boldsymbol{\mathcal{I}}_{\text{adm}} = \mathbf{I}$), problem (9) is in general non-convex. However, if the following additional assumptions are imposed on the scenario,

$$\mathbf{R} = \mathbf{R}_T^T \otimes \mathbf{R}_R, \qquad \mathbf{S} = \mathbf{S}_Q^T \otimes \mathbf{S}_R, \qquad \mathbf{R}_R = \mathbf{S}_R\ , \tag{12}$$

the problem can be diagonalized and the optimal solution can be found in closed form (as a function of a single Lagrange multiplier to be determined), see [7]. In (12), all entities with subscript T, R and Q have dimensions $N_t \times N_t$, $N_r \times N_r$ and $B \times B$, respectively.

[1] $\chi^2_\alpha(n)$ denotes the α percentile of a χ^2 distribution with n degrees of freedom.

The constraint $\mathbf{R}_R = \mathbf{S}_R$ can be relaxed as long as the two matrices share the same eigenvalues, see [8].

The application oriented pilot design formulations of Sect. 2.1 have been analyzed in [11]. Under the assumptions (12) and the additional assumptions $\boldsymbol{\mathcal{I}}_{\text{adm}} = \boldsymbol{\mathcal{I}}_T^T \otimes \boldsymbol{\mathcal{I}}_Q$, it is possible to convert the converse (minimizing $\|\mathbf{P}\|_F^2$ given c) confidence region formulation (11) into a convex problem with closed-form solution. For the average performance formulation (9), yet another assumption $\mathbf{R}_T^{-1} = \boldsymbol{\mathcal{I}}_T$ is needed to turn the problem into diagonalizable form with a closed form solution, up to an unknown permutation of the eigenvalue ordering that has to be found using a combinatorial search.

3. ALTERNATING OPTIMIZATION SOLUTIONS

As has been explained, solutions to the pilot design problems have only been presented under restrictive assumptions on the joint correlation structures of the channel and the interference. Here, we show that a numerical solution can be found easily for generic correlations if we resort to alternating optimization techniques. The trick is to view both $\mathbf{P}$ and the estimation parameters $\mathbf{W}$ as free optimization variables and use the generic MSE matrix expression (4) instead of the concentrated expression (6). The jointly optimal choice of $\mathbf{P}$ and $\mathbf{W}$ will still result in the same solution and therefore alternating optimization (block coordinate descent) will at least find a local optimum of the optimization problem. Convergence is guaranteed, since the cost function is improving in each iteration. The advantage of the alternating optimization approach is that each subproblem is convex and easy to solve for arbitrary $\mathbf{R}$, $\mathbf{S}$ and $\boldsymbol{\mathcal{I}}_{\text{adm}}$. The details are provided below.

Note that these proposed algorithms also apply to the scenario considered in [6, 7, 8], by setting $\boldsymbol{\mathcal{I}}_{\text{adm}} = \mathbf{I}$.

3.1. Average Performance

For a given $\mathbf{P}$, the optimal choice of $\mathbf{W}$ in (9) is (5), see for example [13]. Introduce the notation $\mathbf{p} = \mathrm{vec}\{\mathbf{P}\}$ and let $\mathbf{M}$ be the $N_r^2 N_t B \times N_t B$ selection matrix such that $\mathrm{vec}\{\mathbf{P}^T \otimes \mathbf{I}_{N_r}\} = \mathbf{M}\,\mathrm{vec}\{\mathbf{P}\}$ for all $\mathbf{P} \in \mathbb{C}^{N_t \times B}$. For given $\mathbf{W}$ and using the relationships $\mathrm{Tr}[\mathbf{A}^H\mathbf{B}] = \mathrm{vec}^H\{\mathbf{A}\}\,\mathrm{vec}\{\mathbf{B}\}$ and $\mathrm{vec}\{\mathbf{ABC}\} = (\mathbf{C}^T \otimes \mathbf{A})\,\mathrm{vec}\{\mathbf{B}\}$, the cost function of (9) can then be written as

$$\begin{aligned} \mathrm{Tr}[\boldsymbol{\mathcal{I}}_{\text{adm}} \mathbf{C}_{\text{MSE}}] =& \mathbf{p}^H \mathbf{M}^H (\mathbf{R}^T \otimes \mathbf{W}^H \boldsymbol{\mathcal{I}}_{\text{adm}} \mathbf{W}) \mathbf{M} \mathbf{p}^H \\ &- \mathbf{p}^H \mathbf{M}^H\,\mathrm{vec}\{\mathbf{W}^H \boldsymbol{\mathcal{I}}_{\text{adm}} \mathbf{R}\} \\ &- \mathrm{vec}^H\{\mathbf{W}^H \boldsymbol{\mathcal{I}}_{\text{adm}} \mathbf{R}\} \mathbf{M} \mathbf{p} + \text{const.} \end{aligned}$$

Introducing a Lagrange multiplier λ for the power constraint, it then follows that the optimal $\mathbf{P}$ given $\mathbf{W}$ is given by

$$\begin{aligned} \mathrm{vec}\{\mathbf{P}\} = \Big(\mathbf{M}^H (\mathbf{R}^T \otimes \mathbf{W}^H \boldsymbol{\mathcal{I}}_{\text{adm}} \mathbf{W}) \mathbf{M} + \lambda \mathbf{I}\Big)^{-1} \\ \mathbf{M}^H\,\mathrm{vec}\{\mathbf{W}^H \boldsymbol{\mathcal{I}}_{\text{adm}} \mathbf{R}\}\ , \end{aligned} \tag{13}$$

where bisection or some other line search can be used to find the λ such that $\|\mathbf{P}\|_F^2 = P_{\max}$. It is easy to show that $\|\mathbf{P}\|_F^2$ is monotonically decreasing in λ. It is also easy to show that the optimal λ is in the interval $\Big[0, \|\mathbf{M}^H\,\mathrm{vec}\{\mathbf{W}^H \boldsymbol{\mathcal{I}}_{\text{adm}} \mathbf{R}\}\|/\sqrt{P_{\max}}\Big]$, which therefore can be used as the starting interval for the bisection.

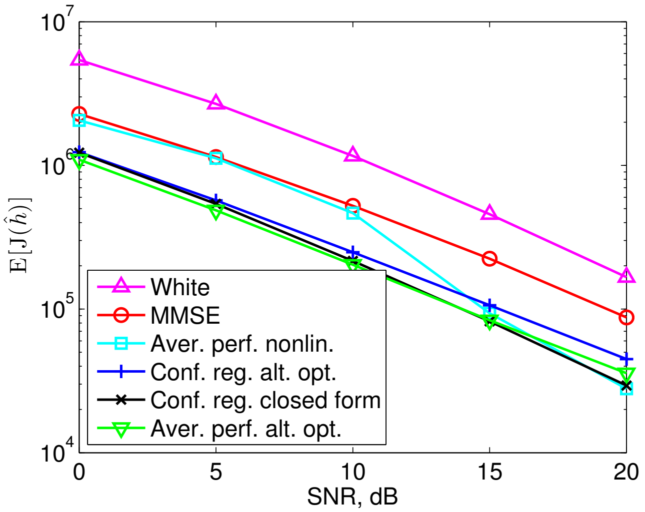

Fig. 1. Average performance in a scenario with known solution to the confidence region problem.

3.2. Using Confidence Regions

For any given $\mathbf{P}$, the MMSE solution (5) minimizes $\mathbf{C}_{\text{MSE}}$ (in the sense of the Löwner order) and is therefore optimal also in (11). For a given $\mathbf{W}$, note that the constraint of (11) is equivalent to $c^{-1}\boldsymbol{\mathcal{I}}_{\text{adm}}^{-1} \succeq \mathbf{C}_{\text{MSE}}$, i.e.

$$c^{-1}\boldsymbol{\mathcal{I}}_{\text{adm}}^{-1} - \mathbf{W}\mathbf{S}\mathbf{W}^H \succeq (\mathbf{W}\tilde{\mathbf{P}} - \mathbf{I})\mathbf{R}(\mathbf{W}\tilde{\mathbf{P}} - \mathbf{I})^H \tag{14}$$

Using a Schur complement and introducing $t = c^{-1}$, we can therefore reformulate (11) as

$$\begin{aligned} \min_{\mathbf{P},t} \quad & t \\ \text{s.t.} \quad & \left[\begin{array}{c|c} t\boldsymbol{\mathcal{I}}_{\text{adm}}^{-1} - \mathbf{W}\mathbf{S}\mathbf{W}^H & \mathbf{W}\tilde{\mathbf{P}} - \mathbf{I} \\ \hline \tilde{\mathbf{P}}^H\mathbf{W}^H - \mathbf{I} & \mathbf{R}^{-1} \end{array}\right] \succeq 0 \\ & \|\mathbf{P}\|_F^2 \leq P_{\max} \;, \end{aligned} \tag{15}$$

which is a convex problem since $\tilde{\mathbf{P}}$ is linear in $\mathbf{P}$ and can be numerically solved using convex solvers like CVX [14, 15].

4. NUMERICAL EXAMPLES

The numerical examples reported here are primarily intended to compare the different algorithmic approaches, but not necessarily to show the end performance of any practical application or propagation scenario. Since the most interesting results are expected in scenarios with reasonably high correlation, all covariance matrices and the matrices $\boldsymbol{\mathcal{I}}_{\text{adm}}$, $\boldsymbol{\mathcal{I}}_T$ and $\boldsymbol{\mathcal{I}}_Q$ where chosen randomly by generating an i.i.d. complex Gaussian matrix $\mathbf{G}$ and forming the Hermitian positive semidefinite matrix as $\mathbf{G}\,\text{diag}\{[1, 2, \ldots, n]\}\mathbf{G}^H$, where n denotes the dimension of the matrix in question. For each generated scenario, $P_{\max}$ was selected to obtain the desired training SNR,

$$\text{SNR} = \frac{P_{\max}\,\text{Tr}\{\mathbf{R}\}}{N_t\,\text{Tr}\{\mathbf{S}\}}$$

In all the experiments, $N_t = 5$, $N_r = 3$ and $B = 6$ were used. The plotted results are averaged over 100 scenario realizations.

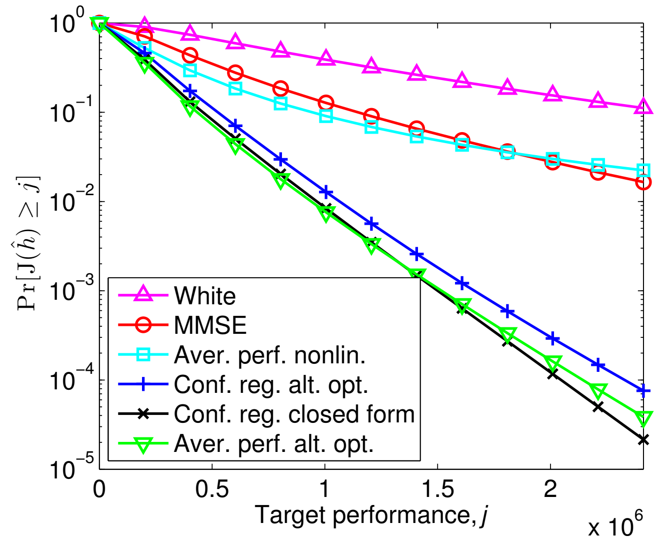

Fig. 2. Outage performance (complementary cumulative distribution function) in a scenario with known solution to the confidence region problem. SNR=10dB.

The following approaches have been compared.

Aver. perf. alt. opt. The proposed alternating optimization approach from Sect. 3.1, using a (single) random initialization.

Conf. reg. alt. opt. The proposed alternating optimization approach from Sect. 3.2, using a (single) random initialization.

Conf. reg. closed form The closed form solution to the converse of (11) from [11], combined with a bisection search over c to obtain the target training power $P_{\max}$. Only applicable under additional assumptions.

Aver. perf. nonlin. Using a standard non-linear optimization procedure to solve (9). The `fmincon` routine in Matlab was used, with the interior point algorithm and a (single) random initialization.

MMSE The closed form solution from [6, 7, 8] optimizing the unweighted MMSE performance, i.e. using $\boldsymbol{\mathcal{I}}_{\text{adm}} = \mathbf{I}$ in (9).

White Using a spatially white $\mathbf{P}$ with columns taken from a scaled FFT matrix.

In figures plotting the outage performance, the exact probabilities have been calculated using Laplace methods (moment generating functions).

4.1. Scenarios with Known Solutions

We first consider scenarios where optimal solutions are known from the literature, namely where (12) holds. We therefore generated random $\mathbf{R}_T$, $\mathbf{R}_R$, $\mathbf{S}_T$, $\boldsymbol{\mathcal{I}}_T$ and $\boldsymbol{\mathcal{I}}_Q$ as described above, form $\mathbf{R}$, $\mathbf{S}$ using (12) and setting $\boldsymbol{\mathcal{I}}_{\text{adm}} = \boldsymbol{\mathcal{I}}_T^T \otimes \boldsymbol{\mathcal{I}}_Q$.

Fig. 1 shows the average performance as a function of SNR. As can be seen, using an optimized pilot gives substantial gain compared to using a standard white pilot or just optimizing the unweighted estimation MMSE. The fact that the closed form solution for the confidence region approach outperforms not only the alternating optimization approach for the same formulation but also that for the average performance (which should give the best result in

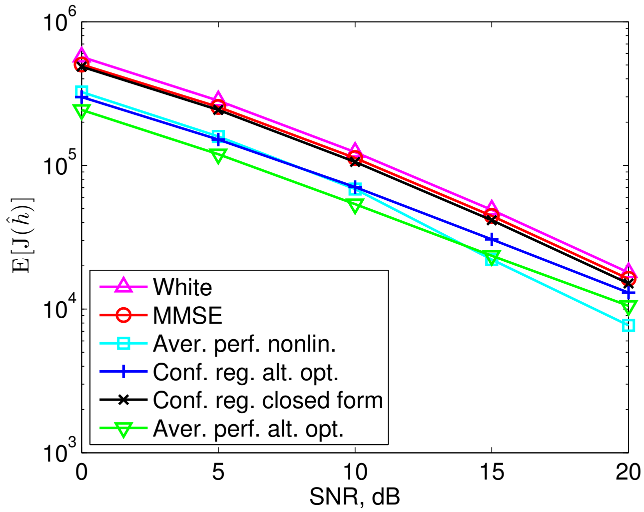

Fig. 3. Average performance in a scenario with unstructured $\mathbf{R}$, $\mathbf{S}$ and $\mathcal{I}_{\text{adm}}$.

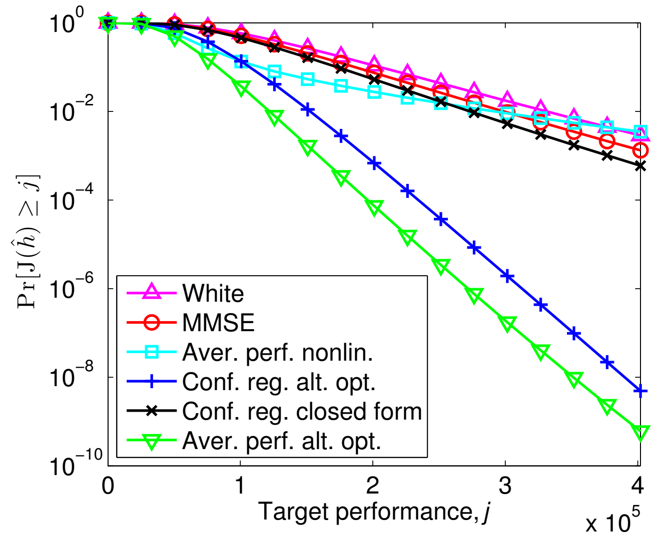

Fig. 4. Outage performance in a scenario with unstructured $\mathbf{R}$, $\mathbf{S}$ and $\mathcal{I}_{\text{adm}}$. SNR=10dB.

this plot if it found the global optimum), indicates that the proposed alternating optimization techniques unfortunately seem to get stuck in local optima. Still, it mostly finds a better local optimum than the generic `fmincon` routine. Using more random initial points is expected to improve the performance, but at the expense of an increased computational complexity. Fig. 2 shows the outage performance for the same methods and scenario. Apart from confirming the conclusions from Fig. 1, it also illustrates the point made in [12] that the confidence region approach does not provide a tight approximation of the outage performance, unless the outage level $1-\alpha$ is extremely small.

4.2. Arbitrarily Correlated Scenarios

As a second example, we consider arbitrary $\mathbf{R}$, $\mathbf{S}$ and $\mathcal{I}_{\text{adm}}$, generated at random as described at the top of Sect. 4. As an ad-hoc engineering solution, we used the strategy of [16] to find approximate Kronecker factorizations minimizing $\|\mathcal{I}_{\text{adm}} - \mathcal{I}_T^T \otimes \mathcal{I}_Q\|_F^2$ and $\|\mathbf{R} - \mathbf{R}_T^T \otimes \mathbf{R}_R\|_F^2 + \|\mathbf{S} - \mathbf{S}_Q^T \otimes \mathbf{R}_R\|_F^2$, respectively. The latter criterion is equivalent to

$$\left\| \begin{bmatrix} \mathbf{R} \\ \mathbf{S} \end{bmatrix} - \begin{bmatrix} \mathbf{R}_T^T \\ \mathbf{S}_Q^T \end{bmatrix} \otimes \mathbf{R}_R \right\|_F^2$$

and can therefore be found using a straightforward generalization of [16]. The resulting Kronecker factors were used as input to the closed form confidence region solution from [11]. Figs. 3 and 4 show that the proposed alternating optimization techniques provide the best performance out of the evaluated techniques, both in terms of average and outage performance. In these examples, there is typically a large mismatch between the true parameters and the Kronecker approximations, which explains why the closed form solution performs so bad.

5. CONCLUSIONS

After a brief review of previous results on optimal training signal design for MIMO channel estimation, we have shown how the resulting optimization problems can be solved without the commonly needed structural constraints on the involved matrices. The proposed method is based on the simple observation that the problems are more tractable if the common trick of concentrating the MSE covariance matrix with respect to the estimation weighting matrix, is *not* used. By keeping this weighting matrix as an optimization parameter, we have shown how all main problem formulations from the literature can be solved using alternating optimization. The numerical examples show that the proposed approach provides good performance in all studied scenarios and that even if they do not always find the global optimum, they often find a better local optimum than using generic non-linear solvers.

6. REFERENCES

[1] P. Stoica and O. Besson, "Training sequence design for frequency offset and frequency-selective channel estimation," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1910–1917, Nov. 2003.

[2] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Trans. IT*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[3] J. Kotecha and A. Sayeed, "Transmit signal design for optimal estimation of correlated MIMO channels," *IEEE Trans. SP*, vol. 52, no. 2, pp. 546–557, Feb. 2004.

[4] M. Biguesh and A. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE Trans. SP*, vol. 54, no. 3, pp. 884–893, Mar. 2006.

[5] T. Wong and B. Park, "Training sequence optimization in MIMO systems with colored interference," *IEEE Trans. Commun.*, vol. 52, no. 11, pp. 1939–1947, Nov. 2004.

[6] D. Katselis, E. Kofidis, and S. Theodoridis, "Training-based estimation of correlated MIMO fading channels in the presence of colored interference," *Sig. Proc.*, vol. 87, pp. 2177–2187, 2007.

[7] Y. Liu, T. Wong, and W. Hager, "Training signal design for estimation of correlated MIMO channels with colored interference," *IEEE Trans. SP*, vol. 55, no. 4, pp. 1486–1497, Apr. 2007.

[8] E. Björnson and B. Ottersten, "A framework for training-based estimation in arbitrarily correlated Rician MIMO channels with Rician disturbance," *IEEE Trans. SP*, vol. 58, no. 3, Mar. 2010.

[9] M. Biguesh, S. Gazor, and M. Shariat, "Optimal training sequence for MIMO wireless systems in colored environments," *IEEE Trans. SP*, vol. 57, no. 8, pp. 3144–3153, Aug. 2009.

[10] H. Hjalmarsson, "System identification of complex and structured systems," *Plenary address European Control Conference / European Journal of Control*, vol. 15, no. 4, pp. 275–310, 2009.

[11] D. Katselis, C. R. Rojas, M. Bengtsson, E. Björnson, X. Bombois, N. Shariati, M. Jansson, and H. Hjalmarsson, "Training sequence design for MIMO channels: An application-oriented approach," *EURASIP Journal on Wireless Communications and Networking*, vol. 2013, no. 245, 2013.

[12] C. R. Rojas, D. Katselis, H. Hjalmarsson, R. Hildebrand, and M. Bengtsson, "Chance constrained input design," in *Proc. IEEE CDC-ECC*, 2011, pp. 2957–2962.

[13] S. Verdú, *Multiuser Detection*, Cambridge University Press, Cambridge, UK, 1998.

[14] I. CVX Research, "CVX: Matlab software for disciplined convex programming, version 2.0," `http://cvxr.com/cvx`, Aug. 2012.

[15] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*, V. Blondel, S. Boyd, and H. Kimura, Eds., Lecture Notes in Control and Information Sciences, pp. 95–110. Springer-Verlag Limited, 2008, `http://stanford.edu/~boyd/graph_dcp.html`.

[16] P. Strobach, "Low-rank detection of multichannel gaussian signals using block matrix approximation," *IEEE Trans. SP*, vol. 43, no. 1, pp. 233–242, Jan. 1995.