A NOVEL SENSING MATRIX DESIGN FOR COMPRESSED SENSING VIA MUTUAL COHERENCE MINIMIZATION

Khaled Ardah‡, Marius Pesavento‡, and Martin Haardt†

‡Communications Research Laboratory (CRL), TU Ilmenau, Ilmenau, Germany
†Communication Systems Group, TU Darmstadt, Darmstadt, Germany
{khaled.ardah, martin.haardt}@tu-ilmenau.de, pesavento@nt.tu-darmstadt.de

ABSTRACT

In compressed sensing, sensing matrices with low mutual coherence should be employed to improve the signal recoverability. To this end, the problem of mutual coherence minimization has attracted attention recently and several methods have been proposed. However, the achievable mutual coherence by the existing methods, in general, is far from the known theoretical lower-bound. In this paper, we propose a novel sensing matrix design method based on mutual coherence minimization, where the original nonconvex problem is relaxed and divided into convex sub-problems, which are updated iteratively using an alternating optimization technique. Simulation results reveal that the proposed method achieves a mutual coherence close to the known lower-bound while outperforming other existing methods.

Index Terms— Compressed sensing, mutual coherence minimization, coordinated beamforming design.

1. INTRODUCTION

In a classical compressed sensing (CS) framework [1–3], the problem is to recover the unknown signal \( x \in \mathbb{C}^K \) from a linear system of measurements given as \( y = Ax \in \mathbb{C}^N \), where \( A \in \mathbb{C}^{N \times K} \) is the so-called sensing matrix. Such a problem appears in many application areas from biomedical signal and image processing [4] to wireless communications [5], and many others, where a closed-form solution can only be achieved using, e.g., the least-squares solutions, when \( N \geq K \). However, when \( N < K \), the system is underdetermined and \( x \) has an infinite number of solutions. Therefore, \( x \) cannot be uniquely recovered without extra assumptions [5, 6]. Nonetheless, if \( x \) has a sufficiently sparse representation and \( A \) is randomly generated, \( x \) can be recovered uniquely from \( y \) with an overwhelming probability using methods such as the basis pursuit (BP) and orthogonal matching pursuit (OMP) [2, 3, 5]. The problem is to find the sparsest solution \( x^* \), i.e., the one that has the smallest number of nonzero components.

The main question here is under which conditions the BP and OMP techniques are guaranteed to recover \( x^* \)? The simplest answer comes from the so-called mutual coherence of the sensing matrix \( A \), denoted as \( \mu(A) \), which is defined as the largest absolute and normalized inner product between the different columns in \( A \) [2]. If \( \mu(A) \) is sufficiently small, the BP and OMP techniques are guaranteed to recover \( x^* \) [2, 5]. More precisely, if \( \|x\|_0 < \frac{1}{2}(1 + \mu^{-1}(A)) \), where \( \|\cdot\|_0 \) denotes the \( \ell_0 \)-norm that counts the nonzero elements. Then the vector \( x \) is the optimal solution, i.e., the sparsest solution and both the BP and OMP techniques are guaranteed to recover it [5]. The goal of this paper is to obtain a sensing matrix with a small mutual coherence.

Although random sensing matrices have an overwhelming recoverability guarantee [2, 3], it has been recently shown that an optimized \( A \), which could be obtained by minimizing \( \mu(A) \), can improve the signal recoverability guarantee [2, 5]. Therefore, the problem of obtaining a sensing matrix with a small mutual coherence has gained an increased attention in the recent years [7–12]. For example, the authors in [7] proposed a gradient-based optimization method with the objective of minimizing the mutual coherence, which consists of one step that is updated iteratively starting from a random initialization. Furthermore, the authors in [8] firstly approximated the non-smooth and nonconvex \( \ell_{\infty} \)-norm by its dual smooth and convex \( \ell_1 \)-norm and then proposed a mutual coherence minimization (MCM) algorithm based on a two-step alternating optimization technique and a matrix projection into the \( \ell_1 \) ball, which is guaranteed to converge, under some conditions, to a local stationary point, although slowly and not monotonically.

In this paper, similarly to [8] and differently from [7, 9, 12], we consider the direct MCM design problem of an \([N \times K]\) sensing matrix, which is a nonconvex and NP-hard problem. To obtain a solution, we first show that the original nonconvex problem can be relaxed and divided it into \( K \) convex
sub-problems. A sensing matrix design method, called hereafter SMCM (sequential mutual coherence minimization), is then proposed based on an alternating optimization technique, where each sub-problem is solved using a sequential projection method originally developed in [13]. Detailed simulation results are presented to evaluate the effectiveness of SMCM, where we show that SMCM outperforms the existing methods in [7,8] and other baseline methods like using random entries and rows of the DFT-matrix.

2. MCM PROBLEM FORMULATION

Using the system model introduced above, we consider a noiseless linear measurement system given as \( \mathbf{y} = \mathbf{Ax} \), where \( \mathbf{y} \in \mathbb{C}^N \) is the measurement vector, \( \mathbf{A} \in \mathbb{C}^{N \times K} \) is the sensing matrix, and \( \mathbf{x} \in \mathbb{C}^K \) is the signal vector of interest with \( \| \mathbf{x} \|_0 \leq s \), meaning that \( \mathbf{x} \) is an \( s \)-sparse vector. The problem is to recover the signal vector \( \mathbf{x} \) while assuming that \( \mathbf{A} \) and \( \mathbf{y} \) are given and \( K \gg N \), i.e., in the underdetermined case. The problem to find the sparsest solution \( \mathbf{x}^* \) can be written as [5]

\[
\mathbf{x}^* = \arg \min_{\mathbf{x}} \| \mathbf{x} \|_0 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{Ax}.
\]

Problem (1) is nonconvex, in general, and the solution requires a combinatorial search among all the possible sparse vectors \( \mathbf{x} \), which is computationally complex and time consuming. Nonetheless, several solution methods have been proposed [2,5] such as the BP technique, which relaxes the nonconvex \( \ell_0 \)-norm to the convex \( \ell_1 \)-norm and the OMP technique, which aims at selecting the best \( s \) vectors from \( \mathbf{A} \) that are mostly correlated with \( \mathbf{y} \) in a greedy manner. However, to guarantee that the above methods succeed in finding \( \mathbf{x}^* \), the perfect reconstruction condition of \( \| \mathbf{x} \|_0 < \frac{1}{2} (1 + \mu^{-1}(\mathbf{A})) \) needs to be satisfied [5], where \( \mu(\mathbf{A}) \) denotes the mutual coherence of the sensing matrix \( \mathbf{A} \) given as

\[
\mu(\mathbf{A}) = \max_{j \neq k} \frac{|\mathbf{a}_j^H \mathbf{a}_k|}{\|\mathbf{a}_j\|_2 \|\mathbf{a}_k\|_2},
\]

with columns \( \mathbf{a}_i \in \mathbb{C}^N \). Here, a large coherence \( \mu(\mathbf{A}) \) means that there exist, at least, two highly correlated columns in \( \mathbf{A} \), which may confuse any pursuit technique, such as BP and OMP. Therefore, to improve recoverability of signal vector \( \mathbf{x} \), \( \mu(\mathbf{A}) \) should be minimized, for which several methods have recently been proposed e.g., in [1,7,8]. In general, the results provided by [1,7,8] confirmed that a well-designed sensing matrix always leads to a better recoverability. However, we noted that the achievable mutual coherence by the aforementioned methods is, in general, far from the known theoretical lower-bound, as we will also show in the numerical results section. Therefore, to improve the signal vector recoverability, an improved sensing matrix design is required.

The problem of sensing matrix design based on MCM can be expressed as [8]

\[
\min_{\mathbf{A} \in \mathbb{C}^{N \times K}} \mu(\tilde{\mathbf{A}}) = \min_{\mathbf{A} \in \mathbb{C}^{N \times K}} f(|\mathbf{G} - \mathbf{I}_K|^2),
\]

where \( \mathbf{G} = \tilde{\mathbf{A}}^H \tilde{\mathbf{A}} \in \mathbb{C}^{K \times K} \) is the so-called Gram-matrix, \( \tilde{\mathbf{A}} = [\mathbf{a}_1 \cdots \mathbf{a}_K] \in \mathbb{C}^{N \times K} \) is a column-wise normalized matrix of \( \mathbf{A} \), \( f(\cdot) \) is a function returning the entry with the largest value, and \( |\cdot| \) is a function returning the element-wise absolute values of the input matrix. Problem (3) is nonconvex and NP-hard [8]. In the following, we show that problem (3) can be relaxed and divided into \( K \) convex sub-problems, which are solved using an alternating optimization technique.

3. PROPOSED MCM METHOD

In this section, we present our method to design a sensing matrix with a low mutual coherence. To start, we first expand matrix \( \mathbf{G}^2 = |\tilde{\mathbf{A}}^H \tilde{\mathbf{A}}|^2 \in \mathbb{C}^{K \times K} \) as

\[
|\mathbf{G}|^2 = \begin{bmatrix}
|\tilde{\mathbf{a}}_1^H \tilde{\mathbf{a}}_1|^2 & \cdots & |\tilde{\mathbf{a}}_1^H \tilde{\mathbf{a}}_K|^2 \\
|\tilde{\mathbf{a}}_2^H \tilde{\mathbf{a}}_1|^2 & \cdots & |\tilde{\mathbf{a}}_2^H \tilde{\mathbf{a}}_K|^2 \\
\vdots & \ddots & \vdots \\
|\tilde{\mathbf{a}}_K^H \tilde{\mathbf{a}}_1|^2 & \cdots & |\tilde{\mathbf{a}}_K^H \tilde{\mathbf{a}}_K|^2
\end{bmatrix} = \begin{bmatrix}
1 & \cdots & |\tilde{\mathbf{a}}_1^H \mathbf{a}_K|^2 \\
\vdots & \ddots & \vdots \\
|\tilde{\mathbf{a}}_K^H \mathbf{a}_1|^2 & \cdots & 1
\end{bmatrix}
\]

As stated in (4), \( |\mathbf{G}|^2 \) is a symmetric matrix with all ones in its main diagonal. Since all vectors in \( \tilde{\mathbf{A}} \) have a unit-norm, we have \( |\mathbf{G}|^2_{i,j} = |\tilde{\mathbf{a}}_i^H \tilde{\mathbf{a}}_j|^2 \leq 1, \forall j \neq k \), where the maximum among them represents the mutual coherence \( \mu(\tilde{\mathbf{A}}) \) of matrix \( \tilde{\mathbf{A}} \). According to [14,15], \( \mu(\tilde{\mathbf{A}}) \) has a theoretical lower-bound given as \( \mu(\tilde{\mathbf{A}}) \geq \sqrt{\beta} \), where \( \beta = \frac{K-N}{N(K-1)} \). This means that, at the best, we have \( \mu(\tilde{\mathbf{A}}) \approx \max_{j \neq k} \{|\tilde{\mathbf{a}}_j^H \tilde{\mathbf{a}}_k|^2\} = \beta \). Noting that each \( k \)th vector \( \tilde{\mathbf{a}}_k \) appears only on the \( k \)th column (and \( k \)th row due to the symmetry of \( |\mathbf{G}|^2 \)), we propose to solve problem (3) in an alternating fashion by iterating over the following \( K \) sub-problems, where the \( k \)th sub-problem for updating the \( k \)th vector \( \tilde{\mathbf{a}}_k \) is given as

\[
\begin{align*}
\text{find} & \quad \tilde{\mathbf{a}}_k, \\
\text{s.t.} & \quad |\tilde{\mathbf{a}}_j^H \tilde{\mathbf{a}}_k|^2 \leq \beta, \forall j \neq k, \\
& \quad |\tilde{\mathbf{a}}_k| = 1,
\end{align*}
\]

for \( k = 1, \ldots, K \). Problems (3) and (5) are related in the sense that both problems are aimed to minimize the maximum off-diagonal entry in (4). However, we note that the strict unit-norm constraint of \( |\tilde{\mathbf{a}}_k| = 1 \) in problem (5) may result in problem infeasibility for poorly initialized vectors \( \tilde{\mathbf{a}}_j, \forall j \neq k \), especially with a tight lower-bound \( \beta \). To avoid such scenario, we propose to relax problem (5) by dropping the unit-norm constraint and only impose it after a solution is obtained, i.e., we first seek a solution to the following relaxed problem

\[
\begin{align*}
\text{find} & \quad \tilde{\mathbf{a}}_k, \\
\text{s.t.} & \quad |\tilde{\mathbf{a}}_j^H \tilde{\mathbf{a}}_k|^2 \leq \beta, \forall j \neq k,
\end{align*}
\]

for \( k = 1, \ldots, K \). Unlike problem (5), problem (6) has a guaranteed feasibility. Note that problem (6) constraints will be
violated after the vector normalization. Nonetheless, our simulation results show that such violation keep decreasing with each iteration, as we will elaborate more about in Remark 2.

To obtain a solution of problem (6), a suitable objective function is needed. One possible approach is as follows

$$\bar{a}_k = \max_{v \in \mathbb{C}^N} |\bar{a}_k^H v|^2,$$

s.t. $|\bar{a}_j^H v|^2 \leq \beta, \forall j \neq k,$

(7)

for $k = 1, \ldots, K$. In problem (7), we borrow the notion from the beamforming design problems in wireless communication systems, see [13], where we interpret $v$ as the beamforming vector that we wish to design so that the desired signal to the $k$th receiver $|\bar{a}_k^H v|^2$ is maximized and the interference signals to the remaining $K-1$ receivers $|\bar{a}_j^H v|^2 \leq \beta, \forall j \neq k$, are minimized for given channel vectors $\{a_1, \ldots, a_K\}$. Then, by exploiting the phase ambiguity, the $k$th sub-problem of (7) can be written equivalently in a convex form as [13]

$$\max_{v \in \mathbb{C}^N} \bar{a}_k^H v,$$

s.t. $|\bar{a}_j^H v|^2 \leq \beta, \forall j \neq k,$

$$\Re(\bar{a}_k^H v) \geq 0 \text{ and } \Im(\bar{a}_k^H v) = 0.$$  

(8)

Note that the constraints $\Re(\bar{a}_k^H v) \geq 0$ and $\Im(\bar{a}_k^H v) = 0$ implies that maximizing $\bar{a}_k^H v$ is equivalent to maximizing $|\bar{a}_j^H v|^2$. Problem (8) is equivalent to [13, problem 3], where the authors showed that the optimal solution to the optimization vector $v$ satisfying the Karush–Kuhn–Tucker conditions [16] of problem (8) has the following form [13, Lemma 1]

$$v^* = \mu_k \bar{a}_k + \sum_{j \in \Gamma_k} \mu_j \bar{a}_j,$$

(9)

for an optimally designed weighting variables $\{\mu_k \in \mathbb{C}\}$ and $\Gamma_k := \{j \neq k : |\bar{a}_j^H v|^2 = \beta\}$. To find $v^*$, the authors proposed a sequential orthogonal projection method, which we summarize in Algorithm 1. The core idea is to initialize the design vector as $v = \bar{a}_k$ and then sequentially project it over the remaining vectors, i.e., $\bar{a}_j, \forall j \neq k$, to satisfy the constraints $|\bar{a}_j^H v|^2 \leq \beta, \forall j \neq k$. Due to the space limitation, we refer to [13] for more details.

Utilizing the above results, Algorithm 2 summarizes the proposed steps to solve problem (8). Starting from an initial sensing matrix $\bar{A}^{(0)}$, at the $r$th outer iteration, Algorithm 2 updates the $k$th column of $\bar{A}^{(r)}$ as $\bar{a}_k^{(r)} = \frac{v}{|v|^2}$, where $v$ is the optimal solution to problem (8) obtained using Algorithm 1. When all the $K$ columns in $\bar{A}^{(r)}$ are updated, Algorithm 2 repeats the same steps at next outer iteration, but now starting from the last designed matrix, i.e., $\bar{A}^{(r)}$.

Remark 1: From all our simulation trails with random initialization of $\bar{A}^{(0)}$, we notice that the cost function $\epsilon = |\mu(\bar{A}^{(r)}) - \mu(\bar{A}^{(r-1)})|^2$ in SMCM always has a monotonic convergence to a point where it stops decreasing or increasing. Nonetheless, this is not a claim that SMCM converges to a local minimum or a stationary point. A convergence proof of SMCM to, at least, a stationary point is left for a follow up research. Moreover, we notice that a different initialization and column-ordering of $\bar{A}^{(0)}$ can lead to a different output. In the numerical section, we show simulations using only random initialization, which can be seen as the worst-case scenario. Finding a better initialization and column-ordering is also left for a future research.

Remark 2: At the final iteration of Algorithm 1, the output vector $v$ satisfies the following equalities and inequalities: $|\bar{a}_j^H v|^2 = \beta$ if $j \in \Phi_k^{(r)}$ and $|\bar{a}_j^H v|^2 \leq \beta$ if $j \notin \Phi_k^{(r)}$, where set $\Phi_k^{(r)}$ collect indices selected in Step 3 [13]. Normalizing the vector $v$ to a unit-norm, the latter equalities and inequalities become $|\bar{a}_j^H v|^2 = \tilde{\beta}$ if $j \in \Phi_k^{(r)}$ and $|\bar{a}_j^H v|^2 \leq \tilde{\beta}$ if $j \notin \Phi_k^{(r)}$, where $\tilde{\beta} = \beta / |v|^2$. Here, we note that $|v|^2$ depends on the input threshold $\beta$, where for a sufficiently small $\beta$, $|v|^2 < 1$ with high probability, which implies that $\tilde{\beta} > \beta$. However, we notice from all our simulation trails that we always have $\mu(\bar{A}^{(0)}) \leq \mu(\bar{A}^{(r+1)})$. A direct implication of this is that SMCM finds in every iteration a new vectors set that have equal or lower maximum cross-correlation than the previously found set. This means that SMCM can increase the vector $v$ norm, while still satisfying problem (8) constraints, i.e., $|v|^2 \rightarrow 1$ and thus $\tilde{\beta} \rightarrow \beta$, while $|\bar{a}_j^H v|^2 \leq \beta, \forall j \in \Phi_k$. 

Remark 3: For the per-iteration cost of Algorithm 1, there are two main parts. The first part is in the projection opera-
tion of the matrix $\Psi$ in Step 1. Since the number of columns of $\Psi$ is increased by one in every $n$ iteration, Step 1 has an overall cost of $O(\sum_{k=0}^{K} (k^3 + nK^2 + kn^2))$, assuming that $N \leq K$. The second part is in the computation of the positive numbers $\mu_1$ in Step 2, which has an overall cost of $O((K-1)N^3)$. Therefore, the overall computational complexity of SMCM is $O(t_1(\sum_{k=0}^{K} (k^3 + nK^2 + kn^2)) + (K-1)N^3))$, where $t_1$ denotes its maximum number of iterations. Meanwhile, the proposed algorithms in [7] and [8] cost $O(t_2N^2K)$ and $O(t_2(N^2K + nK^2))$, where $t_2$ and $t_3$ denotes their maximum number of iterations, respectively.

Remark 4: Consider case when the sensing matrix $\tilde{A}$ is represented as $\tilde{A} = PD$, where $P \in \mathbb{C}^{N \times M}$ is a projection matrix and $D \in \mathbb{C}^{M \times K}$ is a dictionary [7, 8]. We propose to first find $\tilde{A}$ using Algorithm 2 and then obtain the factor matrices by iterating over $P = \tilde{A}D'$ and $D = P'\tilde{A}$ starting from a random initialization. In this case and regardless of the dimensions of the matrices, we can always find $P$ and $D$ such that $\tilde{A} = PD$. In some scenarios, on the other hand, the dictionary matrix $D$ has a specific structure, which is either fixed or can be obtained via training. If $D$ is fixed, then finding $P$ such that $\tilde{A} = PD$ is guaranteed only if $M = K$. Otherwise, the new sensing matrix may incur some performance loss, which increases as the ratio $M/K$ deceases. The problems of finding $P$ for the scenario of a fixed $D$ and $M < K$ and the scenario of a structured $D$ is out of the scope of this paper and we leave it for future research.

4. NUMERICAL RESULTS

In this section, we show some simulation results evaluating the proposed SMCM algorithm as compared against [7, Algorithm 1] and [8, Algorithm 2]. Moreover, we also show simulation results when the sensing matrix is updated randomly from the Gaussian distribution and from the first $N$ rows of the $[K \times K]$ discrete Fourier transform (DFT) matrix. For [8, Algorithm 2], we set $\rho = 0.05$, which acts as an update step-size.

In Fig. 1, we show the achieved mutual coherence versus the number of rows $N$, the number of columns $K$, and the iteration index. From Fig. 1, we can see that the mutual coherence $\mu(\tilde{A})$ of the Random and the DFT matrix approaches are, in general, far from the known lower-bound. On the other hand, we can see that the proposed SMCM method has the best performance, although [8, Algorithm 2] has a comparable performance to the SMCM with a small performance gap. However, unlike [8, Algorithm 2], SMCM has a monotonic and faster convergence rate, within 100 to 200 iterations. Meanwhile, [8, Algorithm 2] convergence depends heavily on the selected parameter $\rho$. Here, we note that [8, Algorithm 2] has a slower convergence rate and small oscillations with a smaller $\rho$ value, and vice-versa otherwise.

On the other hand, although [7, Algorithm 1] has the worst mutual coherence performance, we note that it has the fastest convergence rate, within 1 to 10 iterations, and the lowest computational complexity. Note that, with the exception of the Random method, the mutual coherence $\mu(\tilde{A})$ approaches zeros as the $\frac{N}{K}$ ratio approaches one. In the spacial case of $N = K$, i.e., $\frac{N}{K} = 1$, the optimal mutual coherence of $\mu(\tilde{A}) = 0$ can be simply achieved by the DFT matrix approach. It is worth noting that with a small $\frac{N}{K}$ ratio, the Random method achieves lower mutual coherence as compared to the DFT matrix method, and vice-versa otherwise.

5. CONCLUSIONS

In this paper, we have shown that the nonconvex and NP-hard mutual coherence minimization design problem of a sensing matrix of size $[N \times K]$ can be relaxed and divided into $K$ convex sub-problems, where each sub-problem is solved optimally using existing techniques like the sequential orthogonal projection method. Utilizing this result, a novel mutual coherence minimization method, called SMCM, is proposed based on the alternating optimization technique. Using computer simulations, it is shown that SMCM is capable of obtaining a sensing matrix with a mutual coherence close to the known lower-bound and that it outperforms existing methods.
6. REFERENCES


