# Sensing Matrix Sensitivity to Random Gaussian Perturbations in Compressed Sensing

Anastasia Lavrenko†, Florian Römer†‡, Giovanni Del Galdo†*, and Reiner S. Thomä†*
† *Institute for Information Technology, Technische Universität Ilmenau*
‡ *Fraunhofer Institute for Nondestructive Testing IZFP*
∗ *Fraunhofer Institute for Integrated Circuits IIS*
98693, Ilmenau, Germany
anastasia.lavrenko@tu-ilmenau.de

***Abstract*—In compressed sensing, the choice of the sensing matrix plays a crucial role: it defines the required hardware effort and determines the achievable recovery performance. Recent studies indicate that by optimizing a sensing matrix, one can potentially improve system performance compared to random ensembles. In this work, we analyze the sensitivity of a sensing matrix design to random perturbations, e.g., caused by hardware imperfections, with respect to the total (average) matrix coherence. We derive an exact expression for the average deterioration of the total coherence in the presence of Gaussian perturbations as a function of the perturbations' variance and the sensing matrix itself. We then numerically evaluate the impact it has on the recovery performance.**

***Index Terms*—compressed sensing, sensing matrix, random perturbations, average coherence**

## I. Introduction

Compressed sensing (CS) is a mathematical framework for reduced-rate sampling and recovery of sparse or compressible signals [1], [2]. In its canonical form, it is primarily concerned with estimating an unknown length-$N$ vector $\boldsymbol{x} \in \mathbb{R}^{N\times 1}$ from the following (underdetermined) system of linear equations

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{n}, \tag{1}$$

where $\boldsymbol{y} \in \mathbb{R}^{M\times 1}$ is a length-$M$ vector of observations and $\boldsymbol{A} \in \mathbb{R}^{M\times N}$ is an $M \times N$ sensing matrix. The length-$M$ vector $\boldsymbol{n}$ in (1) represents additive noise, whereas $\boldsymbol{x}$ is assumed to be $K$-sparse, meaning that at most $K \ll N$ of its entries are non-zero. A seminal result in CS states that, under certain conditions on the sensing matrix $\boldsymbol{A}$, $\boldsymbol{x}$ can be recovered from $M < N$ linear observations $\boldsymbol{y}$ [1].

A large body of research is dedicated to studying (1) with a number of powerful recovery algorithms available in the literature [3], [4]. Yet, choosing an appropriate sensing matrix remains a challenge, as a proper $\boldsymbol{A}$ has to ensure the recoverability of $\boldsymbol{x}$ and provide certain performance guarantees in the presence of noise. Traditionally, it is advocated to draw the sensing matrix from random ensembles such as Gaussian or Bernoulli [2]. However, recent results indicate that optimizing a sensing matrix can potentially result in a better recovery performance [5]–[9]. Thus, in [10]–[12] the authors aim at the design that minimizes the so-called matrix coherence. By measuring the largest correlation between the sensing matrix' atoms, smaller matrix coherence yields better worst-case performance guarantees. A similar family of designs considers the degree of correlations between all atoms, the so-called total or average matrix coherence, instead, which allows impoving the average system performance [5], [6], [9], [13], [14].

In this contribution, we investigate the sensitivity of a sensing matrix design to random perturbations in terms of a change of its total coherence. Note that as pointed out in [5], the regular matrix coherence often does not represent the actual behavior of sparse reconstruction algorithms very well. On the other hand, the total mutual coherence is more likely to describe the average CS performance as it provides an average measure of the coherence amoung all dictionary atoms. In practice however, the sensing matrix entries are likely to be subject to a certain degree of alteration, e.g., due to the possible presence of hardware imperfections, phase noise, quantization errors, etc. The discrepancy between the designed sensing matrix and its implemented version is typically characterized by a perturbation matrix that is added to $\boldsymbol{A}$ in (1) [15]. While perturbation analysis in CS context has recently attracted some research attention, it largely focuses on the issues of recovery sensitivity to unknown perturbations [15]–[17] or the design of robust recovery algorithms [18], [19]. Here, we look at this problem from a different perspective: given a certain (arbitrary) sensing matrix we analyze how its total coherence changes in the presence of perturbations. We do so by modeling the perturbations as a random Gaussian process and computing the average total coherence deterioration. The derived exact expression shows that the difference between the total sensing matrix coherence before and after the perturbation depends on the perturbations' variance and the entries of the (unperturbed) sensing matrix itself. We also numerically demonstrate that such perturbations can cause significant performance deterioration, especially when they are unaccounted for. The results of our analysis can be used to mitigate this effect. For instance, one could adapt the design strategy in order to account for a certain degree of variation in a sensing matrix, e.g., by designing $\boldsymbol{A}$ with an additional constraint on the total coherence.

This work was partially supported by the German Research Foundation (DFG) via project CLASS and the Fraunhofer Internal Programs under Grant No. Attract 025-601128.

## II. Total matrix coherence

For any given matrix $\boldsymbol{A}$, we define the total matrix coherence as

$$\epsilon_{\boldsymbol{A}} = \|\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} - \mathbf{I}_N\|_{\mathrm{F}}^2 = \|\boldsymbol{G}_{\boldsymbol{A}} - \mathbf{I}_N\|_{\mathrm{F}}^2, \tag{2}$$

where $\boldsymbol{G}_{\boldsymbol{B}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}$ denotes a Gramian of an arbitrary matrix $\boldsymbol{B}$, while $\|\cdot\|_{\mathrm{F}}$ is the Frobenius norm and $(\cdot)^{\mathrm{T}}$ is the matrix transpose.

Consider now an $M \times N$ perturbation matrix $\boldsymbol{\Delta}$ whose elements are i.i.d. zero-mean Gaussian random variables[1] $\delta_{i,j} \sim \mathcal{N}(0, \sigma^2)$. The perturbed sensing matrix is then expressed as

$$\boldsymbol{A}_{\Delta} = \boldsymbol{A} + \boldsymbol{\Delta}, \tag{3}$$

whereas its total coherence is given by

$$\begin{aligned}\epsilon_{\boldsymbol{A}_{\Delta}} &= \|\boldsymbol{G}_{\boldsymbol{A}_{\Delta}} - \mathbf{I}_N\|_{\mathrm{F}}^2 \\ &= \|(\boldsymbol{A}+\boldsymbol{\Delta})^{\mathrm{T}}(\boldsymbol{A}+\boldsymbol{\Delta}) - \mathbf{I}_N\|_{\mathrm{F}}^2 \\ &= \|\boldsymbol{G}_{\boldsymbol{A}} - \mathbf{I}_N + \boldsymbol{G}_{\boldsymbol{\Delta}} + \boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{A} + \boldsymbol{A}^{\mathrm{T}}\boldsymbol{\Delta}\|_{\mathrm{F}}^2.\end{aligned} \tag{4}$$

For notational convenience, in the following we denote $\boldsymbol{W} = \boldsymbol{G}_{\boldsymbol{A}} - \mathbf{I}_N$ and $\boldsymbol{\Gamma} = \boldsymbol{G}_{\boldsymbol{\Delta}} + \boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{A} + \left(\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{A}\right)^{\mathrm{T}}$ such that

$$\epsilon_{\boldsymbol{A}_{\Delta}} = \|\boldsymbol{W} + \boldsymbol{\Gamma}\|_{\mathrm{F}}^2. \tag{5}$$

Given (2)–(5), the goal of this work is to evaluate the difference between $\epsilon_{\boldsymbol{A}_{\Delta}}$ and $\epsilon_{\boldsymbol{A}}$. Since $\boldsymbol{\Delta}$ is a random matrix, $\epsilon_{\boldsymbol{A}_{\Delta}}$ is a random variable and hence we compute the difference as

$$\epsilon_{\boldsymbol{\Delta}} = \mathrm{E}\{\epsilon_{\boldsymbol{A}_{\Delta}}\} - \epsilon_{\boldsymbol{A}}. \tag{6}$$

Note that that the sensing matrix $\boldsymbol{A}$ is often represented by a product of another $M \times N$ matrix $\boldsymbol{\Phi}$ and some orthogonal dictionary $\boldsymbol{\Psi} \in \mathbb{R}^{N \times N}$, i.e. $\boldsymbol{A} = \boldsymbol{\Phi}\boldsymbol{\Psi}$. Assuming that $\boldsymbol{\Psi}$ is known and taking into account that perturbations will now act on $\boldsymbol{\Phi}$ instead of $\boldsymbol{A}$, (3) transforms into

$$\boldsymbol{A}_{\Delta} = (\boldsymbol{\Phi} + \boldsymbol{\Delta})\boldsymbol{\Psi} = \boldsymbol{\Phi}\boldsymbol{\Psi} + \boldsymbol{\Delta}\boldsymbol{\Psi} = \boldsymbol{A} + \tilde{\boldsymbol{\Delta}}. \tag{7}$$

From the point of view of perturbations' analysis, expression 7 is equivalent to the original formulation (3), but for a possible change of perturbations' variance.

## III. Sensitivity analysis to random perturbations

Taking into account that $\epsilon_{\boldsymbol{A}}$ is deterministic and known, to compute $\epsilon_{\Delta}$ we need to determine the mean value of $\epsilon_{\boldsymbol{A}_{\Delta}}$. We begin by re-writing $\mathrm{E}\{\epsilon_{\boldsymbol{A}_{\Delta}}\}$ as

$$\begin{aligned}\mathrm{E}\{\epsilon_{\boldsymbol{A}_{\Delta}}\} &= \mathrm{E}\{\|\boldsymbol{W}+\boldsymbol{\Gamma}\|_{\mathrm{F}}^2\} = \mathrm{E}\{\mathrm{trace}(\boldsymbol{W}+\boldsymbol{\Gamma})^{\mathrm{T}}(\boldsymbol{W}+\boldsymbol{\Gamma})\} \\ &= \mathrm{trace}(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{W}) + 2\mathrm{E}\{\mathrm{trace}(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{\Gamma})\} + \mathrm{E}\{\mathrm{trace}(\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma})\} \\ &= \epsilon_{\boldsymbol{A}} + 2\mathrm{E}\{\mathrm{trace}(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{\Gamma})\} + \mathrm{E}\{\mathrm{trace}(\boldsymbol{G}_{\boldsymbol{\Gamma}})\} \\ &= \epsilon_{\boldsymbol{A}} + 2\sum_{i=1}^{N} \mathrm{E}\left\{\boldsymbol{w}_i^{\mathrm{T}}\boldsymbol{\gamma}_i\right\} + \sum_{i=1}^{N} \mathrm{E}\left\{\|\gamma_i\|_2^2\right\},\end{aligned} \tag{8}$$

[1]We can often model hardware-induced perturbations, such as the ones caused by phase noise or quantization errors, by a Gaussian process.

where $\boldsymbol{w}_i, \boldsymbol{\gamma}_i$ denote the $i$th columns of $\boldsymbol{W}$ and $\boldsymbol{\Gamma}$, respectively. Inserting (8) into (6) immediately yields

$$\epsilon_{\boldsymbol{\Delta}} = 2\sum_{i=1}^{N} \mathrm{E}\left\{\boldsymbol{w}_i^{\mathrm{T}}\boldsymbol{\gamma}_i\right\} + \sum_{i=1}^{N} \mathrm{E}\left\{\|\gamma_i\|_2^2\right\}. \tag{9}$$

Note that $\boldsymbol{w}_i$ in (8)–(9) is deterministic, while $\boldsymbol{\gamma}_i$ is a random vector with $j$th element composed of

$$\gamma_{i,j} = \boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{\delta}_j + \boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{a}_j + \boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{\delta}_j = \boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{\delta}_j + \beta_{i,j}, \tag{10}$$

where $\beta_{i,j} = \boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{a}_j + \boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{\delta}_j$, while $\boldsymbol{\delta}_i$, $\boldsymbol{a}_i$ denote $i$th columns of $\boldsymbol{\Delta}$, $\boldsymbol{A}$, respectively. Since all $\delta_{i,j}$ are i.i.d Gaussian random variables with equal variance, we have that

$$\beta_{i,j} \sim \mathcal{N}(0, q\alpha_{i,j}^2\sigma^2), \tag{11}$$

where $q = 1 + \delta[i-j]$ and $\alpha_{i,j}^2 = \|\boldsymbol{a}_j\|_2^2 + \|\boldsymbol{a}_i\|_2^2$, while $\delta[n]$ indicates the Kronecker delta function. Given (11), we obtain

$$\mathrm{E}\{\gamma_{i,j}\} = \mathrm{E}\{\boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{\delta}_j\} + \mathrm{E}\{\beta_{i,j}\} = \mathrm{E}\{\boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{\delta}_j\}, \tag{12}$$

where $\boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{\delta}_j$ is the $(i,j)$th element of $\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{\Delta}$ whose mean value and variance are provided by the following theorem.

**Theorem 1.** *Let $\boldsymbol{\Delta}$ be an $M \times N$ random matrix with i.i.d. Gaussian zero-mean elements that have equal variance $\sigma^2$ and $\mathbf{1}_{n\times m}$ denote an all-one matrix of size $n \times m$. Then,*

$$\mathrm{E}\{\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{\Delta}\} = M\sigma^2\mathbf{I}_N \text{ and} \tag{13}$$

$$\mathrm{E}\{(\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{\Delta})^2\} = (M(M+1)\mathbf{I}_N + M\mathbf{1}_{N\times N})\,\sigma^4, \tag{14}$$

*where $(\boldsymbol{B})^n$ means raising the elements of $\boldsymbol{B}$ to $n$th power.*

*Furthermore, denoting by $v_{i,j} = \mathrm{var}\{[\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{\Delta}]_{i,j}\}$ the variance of the $(i,j)$th element of $\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{\Delta}$, we have that*

$$v_{i,j} = qM\sigma^4, \tag{15}$$

*where $q = 1 + \delta[i-j]$.*

*Proof.* Cf. Appendix A. □

Applying Theorem 1, we find that $\mathrm{E}\{\gamma_{i,j}\} = \delta[i-j]M\sigma^2$. Subsequently,

$$\begin{aligned}\sum_{i=1}^{N} \mathrm{E}\left\{\boldsymbol{w}_i^{\mathrm{T}}\boldsymbol{\gamma}_i\right\} &= \sum_{i=1}^{N}\sum_{j=1}^{N} w_{i,j}\mathrm{E}\{\gamma_{i,j}\} = M\sigma^2\sum_{i=1}^{N} w_{i,i} \\ &= M\mathrm{trace}(\boldsymbol{W})\sigma^2 = M\left(\|\boldsymbol{A}\|_{\mathrm{F}}^2 - N\right)\sigma^2.\end{aligned} \tag{16}$$

Now consider the second term of eq. (9). Since $\sum_i \mathrm{E}\left\{\|\gamma_i\|_2^2\right\} = \sum_{i,j}\mathrm{E}\left\{\gamma_{i,j}^2\right\}$, we need to compute $\mathrm{E}\{\gamma_{i,j}^2\}$, which can be written as

$$\begin{aligned}\mathrm{E}\{\gamma_{i,j}^2\} &= \mathrm{E}\{(\boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{\delta}_j + \beta_{i,j})^2\} \\ &= \mathrm{E}\{\beta_{i,j}^2\} + \mathrm{E}\{(\boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{\delta}_j)^2\} + 2\mathrm{E}\{\boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{\delta}_j\beta_{i,j}\}.\end{aligned} \tag{17}$$

To find $\mathrm{E}\{\gamma_{i,j}^2\}$, we examine each term of (17) independently. The value of $\mathrm{E}\{\beta_{i,j}^2\}$ can be found by noticing that

$$\beta_{i,j}^2 \sim \mathrm{Gamma}(0.5, 2q\alpha_{i,j}^2\sigma^2),$$

ISBN 978-90-827970-1-5 © EURASIP 2018

where $\mathrm{Gamma}(k,\theta)$ denotes a Gamma distribution in a shape–scale parametrization. This yields

$$\mathrm{E}\left\{\beta_{i,j}^2\right\} = q\alpha_{i,j}^2\sigma^2. \tag{18}$$

The value of the second term $\mathrm{E}\{(\boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{\delta}_j)^2\}$ is provided by Theorem 1 which states that

$$\mathrm{E}\{(\boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{\delta}_j)^2\} = \begin{cases} M(M+2)\sigma^4, & \text{if } i=j \\ M\sigma^4, & \text{otherwise} \end{cases}. \tag{19}$$

Finally, the last term of (17) can be written as

$$\begin{aligned}\mathrm{E}\{\boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{\delta}_j\beta_{i,j}\} &= \mathrm{E}\{\boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{\delta}_j\boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{a}_j\} + \mathrm{E}\{\boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{\delta}_j(\boldsymbol{a}_i)^{\mathrm{T}}\boldsymbol{\delta}_j\} \\ &= \sum_{m=1}^{M} a_{j,m}\sum_{n=1}^{N}\mathrm{E}\{\delta_{i,m}\delta_{i,n}\}\underbrace{\mathrm{E}\{\delta_{j,n}\}}_{=0} \\ &+ \sum_{m=1}^{M} a_{i,m}\sum_{n=1}^{N}\mathrm{E}\{\delta_{j,m}\delta_{j,n}\}\underbrace{\mathrm{E}\{\delta_{i,n}\}}_{=0} = 0 \end{aligned} \tag{20}$$

Having determined all three terms constituting $\mathrm{E}\{\gamma_{i,j}^2\}$, we can now find $\sum_i \mathrm{E}\left\{\|\gamma_i\|_2^2\right\}$ as

$$\begin{aligned}\sum_i \mathrm{E}\left\{\|\gamma_i\|_2^2\right\} &= \sum_{i,j}\left(q\alpha_{i,j}^2\sigma^2 + \mathrm{E}\{(\boldsymbol{\delta}_i^{\mathrm{T}}\boldsymbol{\delta}_j)^2\}\right) \\ &= 2(N+1)\|\boldsymbol{A}\|_{\mathrm{F}}^2\sigma^2 + NM\left(M+N+1\right)\sigma^4. \end{aligned} \tag{21}$$

At last, we insert (16) and (21) in (9) and finally obtain

$$\begin{aligned}\mathrm{E}\{\epsilon_{\boldsymbol{\Delta}}\} = 2&\left((M+N+1)\|\boldsymbol{A}\|_{\mathrm{F}}^2 - NM\right)\sigma^2 + \\ &+ NM\left(M+N+1\right)\sigma^4. \end{aligned} \tag{22}$$

Expression (22) shows that $\mathrm{E}\{\epsilon_{\boldsymbol{\Delta}}\}$ has a linear and a quadratic term with respect to perturbations' variance $\sigma^2$. It also shows that $\mathrm{E}\{\epsilon_{\boldsymbol{\Delta}}\}$ depends on the entries of the sensing matrix $\boldsymbol{A}$ via $\|\boldsymbol{A}\|_{\mathrm{F}}^2$ as well as the dimensions $M, N$. Interestingly, one can notice that for an arbitrary matrix $\boldsymbol{A}$ with $\|\boldsymbol{A}\|_{\mathrm{F}}^2 < \frac{NM}{N+M+1}$ it is possible to obtain a negative $\mathrm{E}\{\epsilon_{\boldsymbol{\Delta}}\}$ (meaning on average lower total coherence) by adding to it Gaussian perturbations with variance

$$\sigma^2 < 2\left(\frac{1}{M+N+1} - \frac{\|\boldsymbol{A}\|_{\mathrm{F}}^2}{MN}\right).$$

Note however, that this requires that the entries of the sensing matrix $\boldsymbol{A}$ are scaled in a very particular way. Furthermore, in the absence of any additional constraints on $\boldsymbol{A}$, an optimal sensing matrix in terms of the total coherence is known to have $\epsilon_{\boldsymbol{A}_{\mathrm{opt}}} = N - M$ [13], which yields $\|\boldsymbol{A}_{\mathrm{opt}}\|_{\mathrm{F}}^2 = M$. Hence, for a sensing matrix that reaches $\epsilon_{\boldsymbol{A}_{\mathrm{opt}}}$ we have that

$$(M+N+1)\|\boldsymbol{A}_{\mathrm{opt}}\|_{\mathrm{F}}^2 - NM = M(M+1) > 0,$$

and, subsequently, $\mathrm{E}\{\epsilon_{\boldsymbol{\Delta}}\} > 0$, as expected.

## IV. Numerical Evaluation

In this section, we evaluate the influence of Gaussian perturbations on the value of the total matrix coherence numerically. To do so, we first generate a scenario where the elements of $\boldsymbol{A}$ are drawn from a Gaussian distribution such that $a_{i,j} \sim \mathcal{N}\{0, \sigma_{\boldsymbol{A}}^2\}$. In Figure 1, we show the empirical (averaged over $10^4$ realizations of $\boldsymbol{\Delta}$) and the theoretical (both full, according to expression (22), and the two individual terms) difference $\mathrm{E}\{\epsilon_{\boldsymbol{\Delta}}\}/\epsilon_{\boldsymbol{A}}$ against the variance ratio $\sigma_{\boldsymbol{A}}^2/\sigma^2$ for a single realization of $\boldsymbol{A}$ with $N=150$, $M=50$. We can notice that the empirical and theoretical curves coincide, with the quadratic term dominating for large relative perturbations (up to $\sigma^2 = 2(\|\boldsymbol{A}\|_{\mathrm{F}}^2/MN - MN/(M+N+1))$, which corresponds to $\sigma_{\boldsymbol{A}}^2/\sigma^2 \approx -3$ dB in our example) and the linear term dominating for the low level of perturbations.

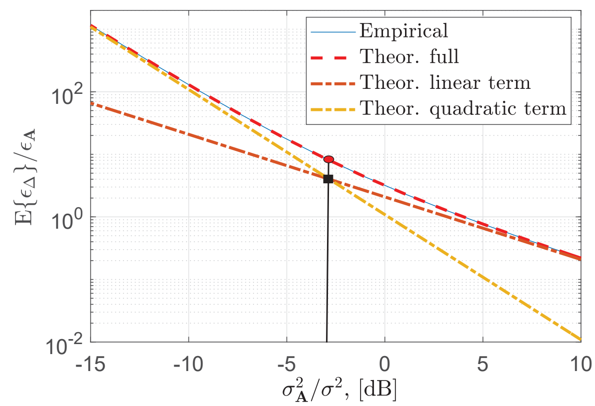


Fig. 1. Empirical and theoretical relative difference $\mathrm{E}\{\epsilon_{\boldsymbol{\Delta}}\}/\epsilon_{\boldsymbol{A}}$ as a function of $\sigma_{\boldsymbol{A}}^2/\sigma^2$ for a single realization of a Gaussian sensing matrix with $N=150$ and $M=50$.

In our next experiment, we change the dimensions $N$, $M$ and evaluate the relative difference $\mathrm{E}\{\epsilon_{\boldsymbol{\Delta}}\}/\epsilon_{\boldsymbol{A}}$ for an optimal (with respect to the total coherence) sensing matrix $\boldsymbol{A}_{\mathrm{opt}}$ with $\epsilon_{\boldsymbol{A}_{\mathrm{opt}}} = N - M$. We depict the resulting $\mathrm{E}\{\epsilon_{\boldsymbol{\Delta}}\}/\epsilon_{\boldsymbol{A}_{\mathrm{opt}}}$ for $\|\boldsymbol{A}_{\mathrm{opt}}\|_{\mathrm{F}}^2/\mathrm{E}\{\|\boldsymbol{\Delta}\|_{\mathrm{F}}^2\} = 1/N\sigma^2 = 0$ dB in Figure 2 in a form of a color plot. We observe that for $N > M$, increasing $N$ with a fixed value of $M$ results in the decrease of the relative difference $\mathrm{E}\{\epsilon_{\boldsymbol{\Delta}}\}/\epsilon_{\boldsymbol{A}_{\mathrm{opt}}}$, while increasing $M$ with a fixed $N$ has an opposite effect.

Finally, we examine how $\mathrm{E}\{\epsilon_{\boldsymbol{\Delta}}\}/\epsilon_{\boldsymbol{A}_{\mathrm{opt}}} \geq 0$ impacts the recovery performance, as judged by the mean squared error (MSE) between the true $\boldsymbol{x}$ and recovered $\hat{\boldsymbol{x}}$ coefficient vectors. Here, to obtain $\hat{\boldsymbol{x}}$ from $\boldsymbol{y} = \boldsymbol{A}_{\boldsymbol{\Delta}}\boldsymbol{x} + \boldsymbol{n}$ we apply the orthogonal matching purcuit (OMP) algorithm [3]. The system dimensions are set to $N=150$ and $M=30$, while the sparsity order is $K=3$. Figure 3 demonstrates normalized MSE, which shows how much the MSE increases due to perturbations compared to the optimal case of $\boldsymbol{A}_{\boldsymbol{\Delta}} = \boldsymbol{A}_{\mathrm{opt}}$, as a function of $\mathrm{E}\{\epsilon_{\boldsymbol{\Delta}}\}/\epsilon_{\boldsymbol{A}_{\mathrm{opt}}}$ for an SNR of 15 dB, where the SNR is defined as $\mathrm{E}\{\|\boldsymbol{A}_{\boldsymbol{\Delta}}\boldsymbol{x}\|_2^2\}/\mathrm{E}\{\|\boldsymbol{n}\|_2^2\}$. Note that the normalized MSE value of 1 means that the MSE provided by the perturbed sensing matrix $\boldsymbol{A}_{\boldsymbol{\Delta}}$ is the same as the one

ISBN 978-90-827970-1-5 © EURASIP 2018

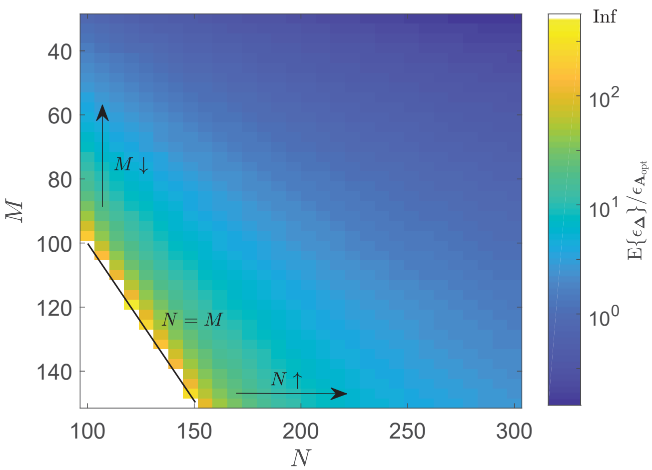


Fig. 2. Relative difference $\mathrm{E}\{\epsilon_\Delta\}/\epsilon_{\boldsymbol{A}_{\mathrm{opt}}}$ as a function of $N$ and $M$ for $\|\boldsymbol{A}_{\mathrm{opt}}\|_{\mathrm{F}}^2/\mathrm{E}\{\|\boldsymbol{\Delta}\|_{\mathrm{F}}^2\} = 1/N\sigma^2 = 0$ dB.

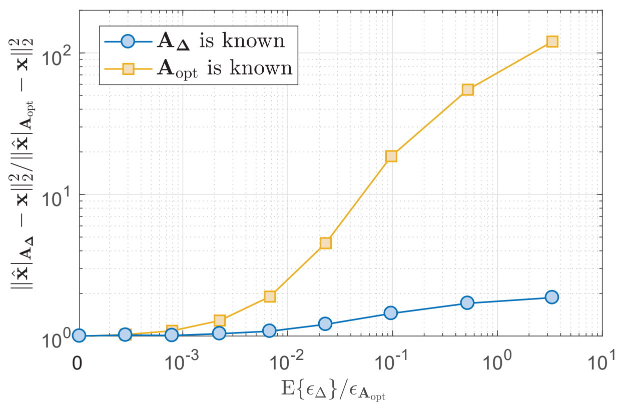


Fig. 3. Normalized MSE as a function of $\mathrm{E}\{\epsilon_{\boldsymbol{\Delta}}\}/\epsilon_{\boldsymbol{A}_{\mathrm{opt}}}$ for $N = 150$, $M = 30$, $K = 3$ and an SNR of 15 dB. The normalized MSE value of 1 means that the MSE provided by the perturbed sensing matrix $\boldsymbol{A}_{\boldsymbol{\Delta}}$ is the same as the one provided by an unperturbed sensing matrix $\boldsymbol{A}_{\mathrm{opt}}$.

provided by $\boldsymbol{A}_{\mathrm{opt}}$. The results are presented for two following cases: $\boldsymbol{x}$ is recovered knowing i) the full perturbed matrix $\boldsymbol{A}_{\boldsymbol{\Delta}} = \boldsymbol{A}_{\mathrm{opt}} + \boldsymbol{\Delta}$ or ii) only $\boldsymbol{A}_{\mathrm{opt}}$. Note that the former can happen when the main source of perturbations is of a deterministic nature, such as quantization errors for instance, which we only model as random, while the latter is the case for perturbation causes such as phase noise. In both cases, we see that in the presence of perturbations the recovery performance deteriorates; its impact is however noticeably more severe when the perturbations are unknown.

## V. Conclusions

In this work, we analyzed the sensitivity of the total coherence of the compressed sensing matrix to random Gaussian perturbations. We derived an exact formula for computing the average total coherence of the perturbed sensing matrix, which shows how it deteriorates in the presence of random perturbations. We then numerically demonstrated the negative effect it has on the recovery performance. Our results support the intuition that for reliable performance the presence of random perturbations should be accounted for either during the sensing matrix design or/and the signal recovery.

## Appendix A. Proof of Theorem 1

Denote by $\delta_{i,j}$ and $g_{p,q}$ the elements of $\boldsymbol{\Delta}$ and $\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{\Delta}$, respectively. To prove Theorem 1, we need to derive the mean values of $g_{p,q}$, $g_{p,q}^2$ and the variance of $g_{p,q}$ for all $p, q = 1, 2, \ldots, N$.

Consider first $g_{p,q}$. For $p \neq q$ we easily obtain

$$\mathrm{E}\{g_{p,q}\} = \mathrm{E}\left\{\sum_{m=1}^{M} \delta_{p,m}\delta_{q,m}\right\} = \sum_{m=1}^{M} \mathrm{E}\{\delta_{p,m}\}\mathrm{E}\{\delta_{q,m}\} = 0$$

due to the independence of $\delta_{p,m}$ and $\delta_{q,m}$. When $p = q$ we have that $g_{p,p} = \|\boldsymbol{\delta}_p\|_2^2$ which means that $g_{p,p}/\sigma^2 \sim \chi^2(M)$, where $\chi^2(k)$ denotes a chi-square distribution with $k$ degrees of freedom. Hence, $g_{p,p} \sim \sigma^2\chi^2(M) = \mathrm{Gamma}(\frac{M}{2}, 2\sigma^2)$, where $\mathrm{Gamma}(k, \theta)$ is the Gamma distribution in a shape–scale parametrization. Therefore,

$$\mathrm{E}\{g_{p,p}\} = \frac{M}{2} 2\sigma^2 = M\sigma^2 \tag{23}$$

for any $p = 1, 2, \ldots, N$. The variance of $g_{p,q}$ can be calculated as follows. For $p = q$ we immediately obtain $\mathrm{var}\{g_{p,p}\} = \frac{M}{2} 4\sigma^4 = 2M\sigma^4$. When $p \neq q$ we have that $g_{p,q} = \sum_{m=1}^{M} \delta_{p,m}\delta_{q,m}$. Let us represent $\delta_{p,m}\delta_{q,m}$ as

$$\delta_{p,m}\delta_{q,m} = \frac{1}{4}(\delta_{p,m} + \delta_{q,m})^2 - \frac{1}{4}(\delta_{p,m} - \delta_{q,m})^2. \tag{24}$$

Since all $\delta_{p,m}$, $\delta_{q,m}$ are i.i.d. zero-mean normal random variables with equal variance $\sigma^2$, we have that $\delta_{p,m} \pm \delta_{q,m}$ is a zero-mean normal random variable with variance $2\sigma^2$. Then,

$$\delta_{p,m}\delta_{q,m} = \frac{1}{2}\sigma^2(Q - R), \tag{25}$$

where $Q, R \sim \chi^2(1)$. Since $\delta_{p,m}$, $\delta_{q,m}$ have the same variance, $Q$ and $R$ are independent and hence

$$\mathrm{var}\{\delta_{p,m}\delta_{q,m}\} = \frac{1}{4} 4\sigma^4 = \sigma^4, \text{ and} \tag{26}$$

$$\mathrm{var}\{g_{p,q}\} = M\sigma^4. \tag{27}$$

For $\mathrm{E}\{g_{p,q}^2\}$ we can now simply write that

$$\begin{aligned}\mathrm{E}\{g_{p,q}^2\} &= \mathrm{var}\{g_{p,q}\} + \mathrm{E}^2\{g_{p,q}\} \\ &= \begin{cases} 2M\sigma^4 + (M\sigma^2)^2, & \text{if } p = q \\ M\sigma^4 + 0, & \text{otherwise} \end{cases}.\end{aligned} \tag{28}$$

ISBN 978-90-827970-1-5 © EURASIP 2018 

## REFERENCES


[1] David L Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[2] E.J. Candes and M.B. Wakin, "An introduction to compressive sampling," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 21–30, 2008.

[3] Joel A Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Transactions on Information theory*, vol. 50, no. 10, pp. 2231–2242, 2004.

[4] Joel A Tropp, "Just relax: Convex programming methods for identifying sparse signals in noise," *IEEE Transactions on Information Theory*, vol. 52, no. 3, pp. 1030–1051, 2006.

[5] Michael Elad, "Optimized projections for compressed sensing," *IEEE Transactions on Signal Processing*, vol. 55, no. 12, pp. 5695–5702, 2007.

[6] Julio Martin Duarte-Carvajalino and Guillermo Sapiro, "Learning to sense sparse signals: Simultaneous sensing matrix and sparsifying dictionary optimization," *IEEE Transactions on Image Processing*, vol. 18, no. 7, pp. 1395–1408, 2009.

[7] Wei Chen, Miguel RD Rodrigues, and Ian J Wassell, "On the design of optimized projections for sensing sparse signals in overcomplete dictionaries," in *IEEE International Conference on Acoustics, Speech and Signal Processing (2012 ICASSP)*, 2012, pp. 3457–3460.

[8] Can-Yi Lu and De-Shuang Huang, "Optimized projections for sparse representation based classification," *Neurocomputing*, vol. 113, pp. 213–219, 2013.

[9] Mohamed Ibrahim, Venkatesh Ramireddy, Anastasia Lavrenko, Jonas König, Florian Römer, Markus Landmann, Marcus Grossmann, Giovanni Del Galdo, and Reiner S Thomä, "Design and analysis of compressive antenna arrays for direction of arrival estimation," *Elsevier Signal Processing*, vol. 138, pp. 35–47, 2017.

[10] Yair Weiss, Hyun Sung Chang, and William T Freeman, "Learning compressed sensing," in *Snowbird Learning Workshop, Allerton, CA*, 2007.

[11] Vahid Abolghasemi, Saideh Ferdowsi, Bahador Makkiabadi, and Saeid Sanei, "On optimization of the measurement matrix for compressive sensing," in *IEEE 2010 18th European Signal Processing Conference*, 2010, pp. 427–431.

[12] Zhouchen Lin, Canyi Lu, and Huan Li, "Optimized projections for compressed sensing via direct mutual coherence minimization," *arXiv preprint*, 2015.

[13] Lihi Zelnik-Manor, Kevin Rosenblum, and Yonina C Eldar, "Sensing matrix optimization for block-sparse decoding," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4300–4312, 2011.

[14] Vahid Abolghasemi, Saideh Ferdowsi, and Saeid Sanei, "A gradient-based alternating minimization approach for optimization of the measurement matrix in compressive sensing," *Signal Processing*, vol. 92, no. 4, pp. 999–1009, 2012.

[15] Matthew A Herman and Thomas Strohmer, "General deviants: An analysis of perturbations in compressed sensing," *IEEE Journal of Selected topics in Signal Processing*, vol. 4, no. 2, pp. 342–349, 2010.

[16] Yuejie Chi, Louis L Scharf, Ali Pezeshki, and A Robert Calderbank, "Sensitivity to basis mismatch in compressed sensing," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2182–2195, 2011.

[17] Zai Yang, Cishen Zhang, and Lihua Xie, "Robustly stable signal recovery in compressed sensing with structured matrix perturbation," *IEEE Transactions on Signal Processing*, vol. 60, no. 9, pp. 4658–4671, 2012.

[18] Oguzhan Teke, Ali Cafer Gurbuz, and Orhan Arikan, "Perturbed orthogonal matching pursuit," *IEEE Transactions on Signal Processing*, vol. 61, no. 24, pp. 6220–6231, 2013.

[19] Reza Arablouei, "Fast reconstruction algorithm for perturbed compressive sensing based on total least-squares and proximal splitting," *Signal Processing*, vol. 130, pp. 57–63, 2017.


ISBN 978-90-827970-1-5 © EURASIP 2018 