A Conjugate Gradient Algorithm for Blind Sensor Calibration in Sparse Recovery

Hao Shen, Martin Kleinsteuber, Cagdas Bilen, Rémi Gribonval

2013 IEEE INTERNATIONAL WORKSHOP ON MACHINE LEARNING FOR SIGNAL PROCESSING, SEPT. 22–25, 2013, SOUTHAMPTON, UK

A CONJUGATE GRADIENT ALGORITHM FOR BLIND SENSOR CALIBRATION IN SPARSE RECOVERY

Hao Shen, Martin Kleinsteuber[*]

Technische Universität München
Arcisstr. 21, 80333 Munich, Germany

Çağdaş Bilen, Remi Gribonval[†]

INRIA, Centre Inria Rennes
Bretagne Atlantique, 35042 Rennes Cedex, France

ABSTRACT

This work studies the problem of blind sensor calibration (BSC) in linear inverse problems, such as compressive sensing. It aims to estimate the unknown complex gains at each sensor, given a set of measurements of some unknown training signals. We assume that the unknown training signals are all sparse. Instead of solving the problem by using convex optimization, we propose a cost function on a suitable manifold, namely, the set of complex diagonal matrices with determinant one. Such a construction can enhance numerical stabilities of the proposed algorithm. By exploring a global parameterization of the manifold, we tackle the BSC problem with a conjugate gradient method. Several numerical experiments are provided to oppose our approach to the solutions given by convex optimization and to demonstrate its performance.

Index Terms— Blind sensor calibration, compressive sensing, conjugate gradient algorithm.

1. INTRODUCTION

The advances in the field of compressive sensing have been one of the important developments in signal processing within the last decade. It has been shown that s-sparse signals can be sampled at much lower rate than required by the Nyquist-Shannon theorem, cf. [1]. More precisely, let $y \in \mathbb{C}^n$ be an s-sparse source vector, $A \in \mathbb{C}^{n\times m}$ be a given measurement matrix with $m \ll n$, the linear measurements of y under A are modeled as

$$z = A^{\mathsf{H}} y + e, \tag{1}$$

where $(\cdot)^{\mathsf{H}}$ denotes the Hermitian transpose, and e is the additive measurement noise. Under certain conditions on the measurement matrix A, the unknown signal y can be recovered accurately by solving the convex optimization problem

$$\operatorname*{argmin}_{y\in\mathbb{C}^n} \|y\|_1, \qquad \text{s.t.} \quad \|z - A^{\mathsf{H}} y\|_2^2 \leq \epsilon. \tag{2}$$

Here $\|\cdot\|_1$ and $\|\cdot\|_2$ denote the ℓ_1-norm and ℓ_2-norm of vectors, respectively, and ϵ is the consistency parameter chosen with respect to the noise power. The ℓ_1-norm is known to favor the selection of sparse signals among the ones satisfying the measurement constraints. Surprisingly, the number of measurements needed for an accurate recovery of y is shown to scale only linearly with the sparsity s, cf. [1].

Unfortunately, in some practical scenarios, it is sometimes not possible to perfectly know the measurement matrix A in advance. For example, in applications dealing with distributed sensors or radars, the location or intrinsic parameters of the sensors are not exactly known. In turn, it results in some unknown phase shifts and/or gains at each sensor, cf. [2, 3]. Similarly, applications with microphone arrays are also shown to require a calibration process of each microphone, in order to account for the unknown gain and phase shifts, cf. [4]. Unlike additive perturbations in the measurement matrix, such a multiplicative perturbation may introduce significant distortion, if it is ignored during recovery, cf. [5, 6].

In order to deal with the multiplicative unknown gains and phase shifts at each sensor, the work in [7, 8] proposes to consider a measurement system with multiple sparse input signals such that the measured signals $\{z_i\}_{i=1}^k$ are modeled as

$$z_i = X A^{\mathsf{H}} y_i + e_i, \tag{3}$$

for all $i = 1, \ldots, k$. Here, the matrix $X \in \mathbb{C}^{m\times m}$ is diagonal, representing the unknown complex valued gains for each sensor. The task of a BSC problem is to recover both the diagonal sensor parameter matrix X and the original sparse source signals $\{y_i\}_{i=1}^k$, given the measurement matrix $A \in \mathbb{C}^{n\times m}$ and the observations $\{z_i\}_{i=1}^k$.

The recovery of $D := X^{-1}$ and y_i up to a global scale in a noise free setting has been shown to be possible with the

[*]This work has been supported by the German DFG funded Cluster of Excellence CoTeSys - Cognition for Technical Systems.

[†]This work was partly funded by the European Research Council, PLEASE project (ERC-StG-2011-277906).

978-1-4799-1180-6/13/$31.00

convex optimization task

$$\underset{D,y_1,\ldots,y_k}{\operatorname{argmin}} \sum_{i=1}^{k} \|y_i\|_1, \quad \text{s.t. } Dz_i = A^{\mathsf{H}} y_i, \quad i = 1,\ldots,k \quad \operatorname{tr}(D) = c \tag{4}$$

with an arbitrary constant $c > 0$. In the presence of noise, however, this approach has a fundamental limitation, since the unknown gains as well as the unknown global scale affects the signal to noise ratio during recovery due to the multiplication of the whole system with the matrix $D = X^{-1}$. Therefore the parameters for the consistency terms in optimization cannot be properly selected without knowing X in advance.

In this paper, we take an alternative optimization method for the measurement system in (3). Specifically, we develop a minimization approach to solve the problem by employing some smooth non-convex sparsity promoting function. In order to enhance numerical stabilities of potential algorithms to minimize the proposed cost function, we propose a numerically stable regularization, i.e. a suitable manifold of the set of complex diagonal matrices with determinant one. Its performance in the current problem setting is investigated both theoretically and numerically.

The rest of the paper is organized as follows. In Section 2, we provide a brief description of the BSC problem, and develop an appropriate regularization. Section 3 constructs a global parameterization of the underlying manifold, which plays an important role in developing the CG algorithm in Section 4. Some numerical experiments are discussed in Section 5, together with a conclusion in Section 6.

2. PROBLEM DESCRIPTIONS AND ITS REGULARIZATIONS

In what follows, we denote by $(\cdot)^\top$ the matrix transpose and by $(\cdot)^*$ the (entry-wise) complex conjugate. Let $z \in \mathbb{C}$, $|z| = \sqrt{zz^*}$ and $\Re z$ denote the modulus and the real part of z, respectively. The complex unit is denoted by $\mathrm{i} := \sqrt{-1}$, and $\log(z)$ computes the principal value of the logarithm of z. We denote by $Gl(m)$ the set of all invertible complex $(m \times m)$-matrices, I_m the $(m \times m)$-identity matrix, $D(m)$ the set of all invertible $(m \times m)$ complex diagonal matrices, i.e.

$$D(m) := \{X \in Gl(m) | X \text{ is diagonal}\}. \tag{5}$$

By $\operatorname{tr}(A)$, $\exp(A)$ and $\|\cdot\|_F$ we denote the trace, the matrix exponential, and the Frobenius norm of matrices, respectively.

Now, let us denote $Y = [y_1,\ldots,y_k] \in \mathbb{C}^{n\times k}$, $Z = [z_1,\ldots,z_k] \in \mathbb{C}^{m\times k}$, and $E = [e_1,\ldots,e_k] \in \mathbb{C}^{m\times k}$. Then the blind sensor calibration model given in (3) is rewritten compactly as

$$Z = XA^{\mathsf{H}}Y + E. \tag{6}$$

By assuming that the additive measuring noise E is zero mean Gaussian, we propose the least squares based cost function

$$\begin{aligned} f\colon D(m) \times \mathbb{C}^{n\times k} &\to \mathbb{R} \\ (X,Y) &\mapsto \tfrac{1}{2}\left\|Z - XA^{\mathsf{H}}Y\right\|_F^2 + \lambda g(Y), \end{aligned} \tag{7}$$

where $g(Y) \geq 0$ is some sparsity promoting penalty function, and the parameter $\lambda > 0$ weighs the sparsity term against the overall residual error.

Minimizing the above cost function f is still an ill-posed problem. There exists the unidentifiable joint scaling ambiguity in the solutions. In order to cope with this situation, the work in [8] proposes to restrict the set of solution matrices to have a constant trace. Here, we follow another approach that is motivated by the sensitivity of solutions of Eq.(3) to noise. Let $b_i = A^{\mathsf{H}} y_i$. Then the noise sensitivity of solutions b_i of $z_i = Xb_i$ is captured by the condition number of X. We argue that a moderate condition number leads to improved numerical stability. To that end, we propose to restrict the diagonal matrix to have determinant one, which leads us to the set

$$Sd(m) := \{X \in D(m) | \det(X) = 1\}. \tag{8}$$

It is a subgroup of the special linear group $Sl(m)$, and its Lie algebra is known to be

$$\mathfrak{sd}(m) := \{X \in \mathbb{C}^{m\times m} | X \text{ is diagonal, } \operatorname{tr}(X) = 0\}. \tag{9}$$

Note that, $Sd(m)$ is a manifold of real dimension $2(m-1)$.

Lemma 1 *Let $X := \operatorname{diag}(x_1,\ldots,x_m) \in Sd(m)$. Then the condition number of X*

$$\operatorname{cond}(X) = \frac{\max_i |x_i|}{\min_i |x_i|} \tag{10}$$

is bounded on the set $\{X \in Sd(m) | \|X\|_F \leq K\}$.

Proof Since $\|X\|_F < K$, we have

$$\max_i |x_i| < K. \tag{11}$$

The fact that $\det(X) = 1$ yields that

$$\min_i |x_i| =: \kappa > 0. \tag{12}$$

Therefore, we get

$$\operatorname{cond}(X) \leq \frac{K}{\kappa}. \tag{13}$$

The result follows. ■

Remark 1 *It is important to notice that, in general for the set $Sd(m)$, the parameter κ can be arbitrarily small, i.e. the condition number of X can be unbounded. However, it is practically reasonable to assume that the largest gain at the sensors is bounded. In other words, the lemma guarantees an upper bounded condition number of X for a realistic measurement system as in* (3)*. Such a mild assumption of a moderate condition number can lead to improved numerical stability. Particularly note, that such bounds do not hold in general for the set of matrices with constant trace $\{X \in D(m) | \operatorname{tr}(X) = c\}$.*

Thus, in the following, we focus on the minimization of

$$\begin{aligned} f_1 \colon Sd(m) \times \mathbb{C}^{n\times k} &\to \mathbb{R} \\ (X, Y) &\mapsto \tfrac{1}{2} \left\| Z - XA^{\mathsf{H}}Y \right\|_F^2 + \lambda g(Y). \end{aligned} \tag{14}$$

Regarding the sparsity measure g, we confine ourselves to the smooth function

$$g \colon \mathbb{C}^{n\times k} \to \mathbb{R}, \quad Y \mapsto \sum_{i=1}^{n}\sum_{j=1}^{k} \log\left(1 + \tfrac{\epsilon}{2}|y_{ij}|^2\right) \tag{15}$$

with $\epsilon > 0$, in order to promote the differentiability of the function f_1.

3. A PARAMETERIZATION OF $Sd(m)$

In this section, we develop a global parameterization of $Sd(m)$. It will be a crucial ingredient to develop the CG algorithm for minimizing f_1.

Lemma 2 *The exponential map*

$$\mu \colon \mathfrak{sd}(m) \to Sd(m), \quad \Omega \mapsto \exp(\Omega) \tag{16}$$

is surjective.

Proof Let $\log(z)$ denote the principle value, i.e. the logarithm of a complex number with imaginary part in $(-\pi, \pi]$. Now let $X := \operatorname{diag}(x_1, \dots, x_m) \in Sd(m)$ be arbitrary. We define

$$\begin{aligned} L := \operatorname{diag}\Big(&\log(x_1), \dots, \log(x_{m-1}), \\ &\log(x_m) - \sum_{i=1}^{m} \log(x_i)\Big). \end{aligned} \tag{17}$$

Then $L \in \mathfrak{sd}(m)$ and $X = \exp(L)$. To see the latter, note that since $\det(X) = 1$, and by using the fact that $\det(\exp(A)) = \exp(\operatorname{tr}(A))$ for any matrix A, we have

$$1 = \exp(\operatorname{tr}(\operatorname{Log}(X))) = \exp\left(\sum_{i=1}^{m} \log(x_i)\right). \tag{18}$$

The result readily follows. ∎

We define an isomorphism from $\mathbb{C}^{m-1}$ to the Lie algebra $\mathfrak{sd}(m)$ as

$$\begin{aligned} \Omega \colon \mathbb{C}^{m-1} &\to \mathfrak{sd}(m), \\ h &\mapsto \operatorname{diag}\Big(h_1, \dots, h_{m-1}, -\sum_{i=1}^{m-1} h_i\Big). \end{aligned} \tag{19}$$

The composition of the maps μ and Ω yields a surjective map $\mu \circ \Omega \colon \mathbb{C}^{m-1} \to Sd(m)$. By a slight abuse of notation, we define the map

$$X \colon \mathbb{C}^{m-1} \to Sd(m), \quad h \mapsto \exp(\Omega(h)). \tag{20}$$

Finally, we end up with the composed cost function

$$\begin{aligned} f_2 \colon \mathbb{C}^{m-1} \times \mathbb{C}^{n\times k} &\to \mathbb{R}, \\ (h, Y) &\mapsto f_1(X(h), Y). \end{aligned} \tag{21}$$

4. A CONJUGATE GRADIENT ALGORITHM

In this section, we construct a conjugate gradient (CG) algorithm for minimizing the cost function f_2. CG algorithms are well known for its superlinear rate of convergence and the applicability to large scale optimization problems with low computational complexity, e.g. in sparse recovery [9].

Firstly, we compute the first derivative of f_2 at (h, Y) in the direction $(\xi, \Phi) \in \mathbb{C}^{m-1} \times \mathbb{C}^{n\times k}$ as

$$\begin{aligned} \mathrm{D}f_2(h, Y)(\xi, \Phi) = &\sum_{i=1}^{m-1} \Re\xi_i^*(r_{ii} - r_{mm}) + \\ &+ \Re \operatorname{tr}\big(\Phi^{\mathsf{H}} A(X(h))^{\mathsf{H}}(X(h)A^{\mathsf{H}}Y - Z)\big) + \\ &+ \lambda \sum_{i=1}^{n}\sum_{j=1}^{k} \frac{\epsilon\, \Re\phi_{ij}^* y_{ij}}{1 + \frac{\epsilon}{2}|y_{ij}|^2}, \end{aligned} \tag{22}$$

with $R = (r_{ij}) := (X(h)A^{\mathsf{H}}Y - Z)Y^{\mathsf{H}}A(X(h))^{\mathsf{H}} \in \mathbb{C}^{m\times m}$. It is straightforward to compute the Euclidean gradient of f_2 with respect to the two parameters as

$$\nabla_{f_2}(h) = \left[r_{11} - r_{mm}, \dots, r_{(m-1)(m-1)} - r_{mm}\right]^{\top} \tag{23}$$

and

$$\nabla_{f_2}(Y) = A(X(h))^{\mathsf{H}}\left(X(h)A^{\mathsf{H}}Y - Z\right) + \lambda\nabla_g(Y) \tag{24}$$

with

$$\nabla_g(Y) = \left(\frac{\epsilon\, y_{ij}}{1 + \frac{\epsilon}{2}|y_{ij}|^2}\right)_{ij}. \tag{25}$$

Denoting by $G := (\nabla_{f_2}(h), \nabla_{f_2}(Y)) \in \mathbb{C}^{m-1} \times \mathbb{C}^{n\times k}$ the gradient of f_2 and $H \in \mathbb{C}^{m-1} \times \mathbb{C}^{n\times k}$ the conjugate gradient direction, we present a conjugate gradient algorithm for minimizing the function f_2 as defined in (21), as in Algorithm 1. Step 4 requires to find the local or global minimum

Algorithm 1: A CG blind sensor calibration algorithm

Input : $A \in \mathbb{C}^{n\times m}$ and $Z \in \mathbb{C}^{m\times k}$;

Output: $X(h) \in Sd(m)$ and $Y \in \mathbb{C}^{n\times k}$;

Step 1: Generate initialization $h^{(0)} \in \mathbb{C}^{m-1}$ and $Y^{(0)} \in \mathbb{C}^{n\times k}$, and set $j = 1$;

Step 2: Compute $G^{(1)} = H^{(1)} \leftarrow -(\nabla_{f_2}(h^{(0)}), \nabla_{f_2}(Y^{(0)}))$ using Eq. (23) and Eq. (24) ;

Step 3: Set $j = j + 1$;

Step 4: Update $(h^{(j+1)}, Y^{(j+1)}) \leftarrow (h^{(j+1)}, Y^{(j+1)}) + \lambda H^{(j)}$, where λ is computed as in Eq. (26) ;

Step 5: Update $H^{(j+1)} \leftarrow -G^{(j+1)} + \gamma H^{(j)}$, where $G^{(j+1)} = (\nabla_{f_2}(h^{(j+1)}), \nabla_{f_2}(Y^{(j+1)}))$, and γ is chosen according to Eq. (29) ;

Step 6: If $j \bmod 2(nk+m) - 3 = 0$, set $H^{(j+1)} \leftarrow -G^{(j+1)}$;

Step 7: If $\left\|G^{(j+1)}\right\|$ is small enough, stop. Otherwise, go to Step 3;

of a restricted cost function given the conjugate search directions, which is often unfeasible in practice. In this work, we employ a one-dimensional Newton step instead, i.e.

$$\lambda = -\frac{\frac{\mathrm{d}}{\mathrm{d}\,t} f_2(h+t\xi, Y+t\Phi)\big|_{t=0}}{\frac{\mathrm{d}^2}{\mathrm{d}\,t^2} f_2(h+t\xi, Y+t\Phi))\big|_{t=0}}, \tag{26}$$

where the numerator and the denominator are computed as

$$\begin{aligned}
&\tfrac{\mathrm{d}}{\mathrm{d}\,t} f_2(h+t\xi, Y+t\Phi)\big|_{t=0} \\
&= \Re\,\mathrm{tr}\Big((\Omega(\xi))^{\mathsf{H}}(X(h)A^{\mathsf{H}}Y - Z)Y^{\mathsf{H}}A(X(h))^{\mathsf{H}} \\
&\qquad + \Phi^{\mathsf{H}}A(X(h))^{\mathsf{H}}\left(X(h)A^{\mathsf{H}}Y - Z\right)\Big) \\
&\quad + \lambda \sum_{i=1}^{n}\sum_{j=1}^{k} \frac{\epsilon\,\Re\phi_{ij}^* y_{ij}}{1+\frac{\epsilon}{2}|y_{ij}|^2},
\end{aligned} \tag{27}$$

and

$$\begin{aligned}
&\tfrac{\mathrm{d}^2}{\mathrm{d}\,t^2} f_2(h+t\xi, Y+t\Phi)\big|_{t=0} \\
&= \Re\,\mathrm{tr}\Big((\Omega(\xi)^2)^{\mathsf{H}}(X(h)A^{\mathsf{H}}Y - Z)Y^{\mathsf{H}}A(X(h))^{\mathsf{H}} \\
&\qquad + (\Omega(\xi))^{\mathsf{H}}\Omega(\xi)X(h)A^{\mathsf{H}}YY^{\mathsf{H}}A(X(h))^{\mathsf{H}} \\
&\qquad + 2\Phi^{\mathsf{H}}A(X(h))^{\mathsf{H}}(\Omega(\xi))^{\mathsf{H}}(X(h)A^{\mathsf{H}}Y - Z) \\
&\qquad + 2\Phi^{\mathsf{H}}A(X(h))^{\mathsf{H}}X(h)\Omega(\xi)A^{\mathsf{H}}Y \\
&\qquad + \Phi^{\mathsf{H}}A(X(h))^{\mathsf{H}}X(h)A^{\mathsf{H}}\Phi\Big) \\
&\quad + \lambda \sum_{i=1}^{n}\sum_{j=1}^{k} \frac{\epsilon|\phi_{ij}|^2}{1+\frac{\epsilon}{2}|y_{ij}|^2} - \left(\frac{\epsilon\,\Re\phi_{ij}^* y_{ij}}{1+\frac{\epsilon}{2}|y_{ij}|^2}\right)^2.
\end{aligned} \tag{28}$$

For updating the direction parameter γ in Step 5, we confine ourselves to a formula, which is proposed in [10, 11], as

$$\gamma = \frac{\left\langle G^{(j+1)}, G^{(j+1)} - G^{(j)}\right\rangle}{\left\langle H^{(j)}, G^{(j)}\right\rangle}, \tag{29}$$

where

$$\langle (h_1, Y_1), (h_2, Y_2)\rangle = \Re h_1^{\mathsf{H}} h_2 + \Re\,\mathrm{tr}(Y_1^{\mathsf{H}} Y_2). \tag{30}$$

Finally in Step 6, the search direction is periodically reset to the negative of the gradient, in order to achieve fast convergence.

5. NUMERICAL EXPERIMENTS

In order to test the performance of the proposed algorithm, we set the signal size to $n = 100$. All the non-zero entries in the input signals y_i are randomly generated from an i.i.d. normal distribution, and their positions are chosen uniformly at random in $\{1, \ldots, n\}$. The magnitude of the gains are generated using the normal distribution, i.e. $|d_i| \sim \mathcal{N}(0, 1)$, and the phase of the gains are chosen uniformly at random from the interval $\left[0, \frac{4}{3}\pi\right)$. Note, that the choice of this internal determines the scale of ambiguity in the phases. In particular, the maximum possible ambiguity is observed at $[0, 2\pi)$.

We compare the performance of the CG algorithm in terms of perfect reconstruction, to the solutions given by a convex recovery approach proposed in [8]. The perfect reconstruction is measured by the absolute correlation factor ψ, which is defined as

$$\psi(Y, \widetilde{Y}) := \frac{1}{k}\sum_{i=1}^{k} \frac{|y_i^{\mathsf{H}}\widetilde{y}_i|}{\|y_i\|_2\|\widetilde{y}_i\|_2}, \tag{31}$$

where $Y := [y_1, \ldots, y_k]$ is the ground truth training signals and $\widetilde{Y} := [\widetilde{y}_1, \ldots, \widetilde{y}_k]$ is the corresponding estimates. In our experiments, we take the solutions from the convex recovery approach as a warm starting for the CG algorithm. We set the number of measurements $m = \{40, 50, 60, 70\}$ and the corresponding sparsities $s = \{8, 10, 12, 14\}$, and run the experiments 100 times. The quartile based boxplot of the absolute correlation factor for both methods are drawn in Figure 1 and Figure 2. Our proposed CG approach outperforms consistently the convex approach in terms of average performance. In particular, performance of the proposed CG algorithm, shown in Figure 2, is significantly improved with an increasing number of measurements.

6. CONCLUSIONS

In this paper, we study the problem of blind sensor calibration, propose a numerical stable regularization, and develop an intrinsic conjugate gradient algorithm. Our experimental

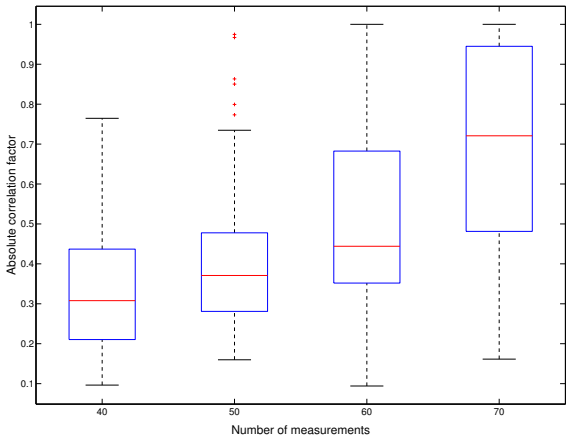

Fig. 1. Absolute correlation factor given by the convex approach.

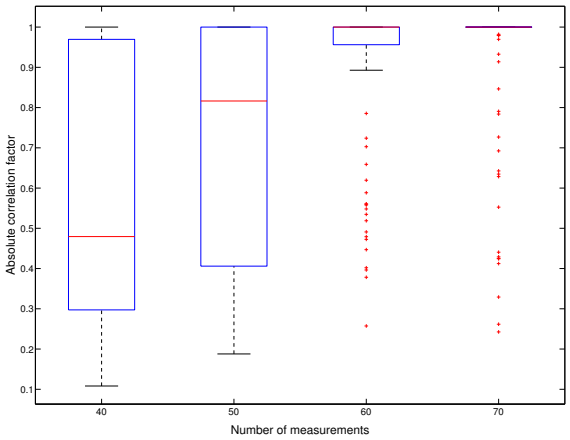

Fig. 2. Absolute correlation factor given by the proposed CG approach.

results have demonstrated its promising performance compared to the convex recovery approach. One limit of the current work is that the performance of the proposed CG algorithm is very sensitive to the initialization process. Specifically, solutions given by the convex recovery approach are taken as warm starts for the CG algorithm. Our future work is to employ more robust line search strategies in the CG framework, in order to reach higher chance of finding a global minimum as well as to increase the convergence speed. Furthermore, a thorough study of the phase transition in terms of the probability of recovery is a work in progress.

7. REFERENCES

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

[2] Z. Yang, C. Zhang, and L. 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.

[3] Boon Chong Ng and Chong-Meng Samson See, "Sensor-array calibration using a maximum-likelihood approach," *IEEE Transactions on Antennas and Propagation*, vol. 44, no. 6, pp. 827–835, 1996.

[4] R. Mignot, L. Daudet, and F. Ollivier, "Compressed sensing for acoustic response reconstruction: Interpolation of the early part," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, 2011, pp. 225–228.

[5] E. J. Candés, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.

[6] M. A. Herman and T. 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.

[7] R. Gribonval, G. Chardon, and L. Daudet, "Blind calibration for compressed sensing by convex optimization," in *Proceedings of the* 37^{th} *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2012, pp. 2713–2716.

[8] Ç. Bilen, G. Puy, R. Gribonval, and L. Daudet, "Blind sensor calibration in sparse recovery using convex optimization," in *the* 10^{th} *International Conference on Sampling Theory and Applications (SAMPTA)*, 2013.

[9] S. Hawe, M. Kleinsteuber, and K. Diepold, "Cartoon-like image reconstruction via constrained ℓ_p-minimization," in *Proceedings of the* 37^{th} *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2012, pp. 717–720.

[10] Y. Liu and C. Storey, "Efficient generalized conjugate gradient algorithms, part 1: Theory,," *Journal of Optimization Theory and Applications*, vol. 69, no. 1, pp. 129–137, 1991.

[11] M. Kleinsteuber and K. Hüper, "An intrinsic CG algorithm for computing dominant subspaces," in *Proceedings of the* 32^{nd} *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2007, pp. IV1405–IV1408.