

BLIND PHASE CALIBRATION IN SPARSE RECOVERY

Çağdaş Bilen* Gilles Puy† Rémi Gribonval* Laurent Daudet‡

* INRIA, Centre Inria Rennes - Bretagne Atlantique, 35042 Rennes Cedex, France.

† Institute of Electrical Engineering, Ecole Polytechnique Federale de Lausanne (EPFL)
CH-1015 Lausanne, Switzerland.

‡ Institut Langevin, CNRS UMR 7587, UPMC, Univ. Paris Diderot, ESPCI, 75005 Paris, France.

ABSTRACT

We consider a *blind* calibration problem in a compressed sensing measurement system in which each sensor introduces an unknown phase shift to be determined. We show that this problem can be approached similarly to the problem of phase retrieval from quadratic measurements. Furthermore, when dealing with measurements generated from *multiple* unknown (but sparse) signals, we extend the approach for phase retrieval to solve the calibration problem in order to recover the signals *jointly* along with the phase shift parameters. Additionally, we propose an alternative optimization method with less computation complexity and memory requirements. The proposed methods are shown to have significantly better recovery performance than individual recovery of the input signals when the number of input signals is sufficiently large.

Index Terms— Compressed sensing, calibration, phase estimation, phase retrieval, lifting

1. INTRODUCTION

Compressed sensing theory shows that K -sparse signals can be sampled at much lower rate than apparently required by the Nyquist-Shannon theorem [1]. More precisely, if $\mathbf{x} \in \mathbb{C}^N$ is a K -sparse source vector then it can be captured by collecting only $M \ll N$ linear measurements

$$y_i = \mathbf{m}_i' \mathbf{x}, \quad i = 1, \dots, M \quad (1)$$

In the above equation, $\mathbf{m}_1, \dots, \mathbf{m}_M \in \mathbb{C}^N$ are *known* measurement vectors, and $'$ denotes the conjugate transpose operator. Under certain conditions on the measurement vectors, the signal can be accurately reconstructed by solving, e.g.,

$$\mathbf{x}_{\ell_1}^* = \arg \min_{\mathbf{z}} \|\mathbf{z}\|_1 \quad \text{subject to} \quad y_i = \mathbf{m}_i' \mathbf{z}, \quad i = 1, \dots, M$$

where $\|\cdot\|_1$ denotes the ℓ_1 -norm, which favors the selection of sparse signals among the ones satisfying the measurement

This work was partly funded by the Agence Nationale de la Recherche (ANR), project ECHANGE (ANR-08-EMER-006) and by the European Research Council, PLEASE project (ERC-StG-2011-277906). LD is on a joint affiliation between Univ. Paris Diderot and Institut Universitaire de France.

constraints. It has been showed that the number of measurements needed for accurate recovery of \mathbf{x} scales only linearly with K [1].

In some practical situations it is not possible to perfectly know the measurement vectors $\mathbf{m}_1, \dots, \mathbf{m}_M$. In applications with sensors of uncertain locations or intrinsic parameters such as radio imaging equipment or microphone arrays, the measurement vectors have an unknown phase shift or scale which severely affects the reconstruction quality if ignored. We call the problem of signal recovery and determination of unknown parameters in this perturbed system without any special calibration input as *the compressive calibration problem*, and a simplified version of this problem dealing with only unknown phase shifts as *the phase calibration problem*. An extension of compressed sensing recovery which estimates the unknown gains along with the unknown sparse signals has been presented in [2], and it is shown that the recovery fails when the sensors have unknown random phase shifts. Therefore an alternative approach to handle such cases is needed.

A completely different but relevant problem arises for applications such as optical interferometric imaging for astronomy in which one has only access to the magnitude of the measurements $z_i = |y_i|^2 = \mathbf{m}_i' \mathbf{x} \mathbf{x}' \mathbf{m}_i$, $i = 1, \dots, M$. Reconstructing the original signal from such magnitude measurements is called the phase retrieval problem which involves similar challenges to phase calibration. Nevertheless, Candès *et al.* have recently showed [3] that \mathbf{x} could be recovered exactly by solving a convex optimization problem with the number of measurements, $M > N$, essentially proportional to N . The method relies on finding a positive semi-definite matrix $\mathbf{X} \triangleq \mathbf{x} \mathbf{x}'$ of rank-one such that $|y_i|^2 = \mathbf{m}_i' \mathbf{X} \mathbf{m}_i$, $i = 1, \dots, M$. The following convex optimization problem can then be solved to recover \mathbf{X} :

$$\begin{aligned} \mathbf{X}^* = \arg \min_{\mathbf{Z}} \quad & \text{trace}(\mathbf{Z}) \\ \text{subject to} \quad & \mathbf{Z} \succcurlyeq 0 \\ & |y_i|^2 = \mathbf{m}_i' \mathbf{Z} \mathbf{m}_i, \quad i = 1, \dots, M \end{aligned} \quad (2)$$

The trace norm $\text{trace}(\cdot)$ favors the selection of low rank ma-

trices among all the ones satisfying the constraints. Let us acknowledge that this problem was also previously studied theoretically in, e.g., [4], but a larger number of measurements is needed for fast reconstruction of the original signal with the technique therein (on the order of N^2 instead of N). Note also that several simple iterative algorithms such as the one described in [5] have been proposed to estimate the signal \mathbf{x} from magnitude measurements, however there is in general no guarantee that such algorithms converge.

When the measured vector \mathbf{x} is sparse, a modification of this so-called Phaselift approach was then proposed by Ohlsson *et al.* [6, 7]. This new approach is called Compressive Phase Retrieval via Lifting (CPRL) and the optimization method is named as Quadratic Basis Pursuit which consists in solving the problem (2) with the addition of a cost term that penalizes non-sparse matrices.

$$\begin{aligned} \mathbf{X}^* &= \arg \min_{\mathbf{Z}} \text{trace}(\mathbf{Z}) + \lambda \|\mathbf{Z}\|_1 & (3) \\ \text{subject to } \mathbf{Z} &\succeq 0 \\ |y_i|^2 &= \mathbf{m}'_i \mathbf{Z} \mathbf{m}_i, \quad i = 1, \dots, M, \end{aligned}$$

where $\lambda > 0$. The authors also provide bounds for guaranteed recovery of this method using a generalization of restricted isometry property.

In this paper, we study the phase calibration problem where the measurements y_1, \dots, y_M are accessible but the measurement vectors $\mathbf{m}_1, \dots, \mathbf{m}_M$ are not precisely known. Unlike earlier studies on uncertain measurement matrices [8], we restrict our study to the case where each sensor introduces an unknown, but constant, phase shift on the measurements. To calibrate the system, we propose to measure a few, but unknown, sparse signals via this not perfectly known sensing system. Then, we show that it is possible to reconstruct these signals via a similar approach to the compressive phase retrieval method (3) used in phase retrieval. Differently from CPRL however, the proposed method provides joint recovery of the input signals and the number of necessary input signals as well as the sparsity play a significant role to determine whether perfect recovery is possible. The provided experimental results show that it is possible to significantly improve upon the recovery performance of CPRL given that the number of signals is sufficiently large. We also propose an alternative recovery method that is more scalable in terms of number of input signals and has less computational complexity as well as less memory requirements than the direct extension of the phase retrieval approach.

2. PROBLEM FORMULATION

Suppose that the measurement system in (1) is perturbed by complex gains at each sensor i and there are multiple sparse input signals, $\mathbf{x}_l \in \mathbb{C}^N$, $l = 1 \dots L$, applied to the system

such that

$$y_{i,l} = d_i e^{j\theta_i} \mathbf{m}'_i \mathbf{x}_l \quad i = 1 \dots M, \theta_i \in [0, 2\pi), d_i \in \mathbb{R}^+ \quad (4)$$

This problem can be studied with two special cases:

- *Gain (magnitude) Calibration:* The phase shifts, θ_i , are known and calibration consists of determining the unknown real gains, $d_i \in \mathbb{R}^+$. This problem has been studied in [2] and it is shown that recovery of d_i and \mathbf{x}_l is possible with convex relaxation.
- *Phase Calibration:* The gain magnitudes, d_i , are known, and calibration consists in determining the unknown phase shifts for each sensor, θ_i .

In this paper we consider the latter scenario, hence $d_i \mathbf{m}_i$ is simply replaced with \mathbf{m}_i for the rest of the discussions. We focus only on the noiseless case for the sake of simplicity.

2.1. Fundamental differences with CPRL

Unlike the classic phase retrieval problem, in phase calibration we have access to the amplitude and the phase information of each measurement $y_{i,l}$. However the phase shifts θ_i , $i = 1 \dots M$, at the sensors are unknown, preventing us to use standard compressed sensing algorithm to reconstruct the training signals. When $L = 1$, these arbitrary phase shifts cannot be determined, and only the magnitude of the measurements can be used to reconstruct the input signal. The case $L = 1$ is thus equivalent to the phase retrieval problem discussed in [3, 6, 9]. However, the problem becomes *fundamentally* different when $L > 1$. Making use of the phase information and noticing that the measurements $y_{i,l}$, $l = 1 \dots L$, for sensor i are *correlated* through the *same* phase shift factor, we propose two methods exploiting these correlations and phase information to reach better recovery performance than one could obtain using CPRL and disregarding all the measured phase.

Furthermore, in comparison to CPRL, a new interesting question arises. While, for CPRL, one is interested in how many measurements are sufficient to recover the initial sparse signal, we are also interested in the question of how many training signals are sufficient to calibrate the system. Even though the optimization problem to solve for calibration is also much more complex compared to CPRL, a solution to handle this issue is proposed in the following sections.

Finally, the potential applications of the proposed method and phase retrieval algorithms are also significantly different as discussed earlier.

2.2. Joint Phase Calibration

Let us suppose that we have in hand uncalibrated measurements $y_{i,k}$ from $L > 1$ training signals. We define the joint

signal matrix $\mathbf{X} \in \mathbb{C}^{LN \times LN}$ as

$$\mathbf{X} \triangleq \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_L \end{bmatrix} [\mathbf{x}'_1 \cdots \mathbf{x}'_L] = \mathbf{x}\mathbf{x}' = \begin{bmatrix} \mathbf{X}_{1,1} & \cdots & \mathbf{X}_{1,L} \\ \vdots & \ddots & \vdots \\ \mathbf{X}_{L,1} & \cdots & \mathbf{X}_{L,L} \end{bmatrix} \quad (5)$$

which is rank-one, positive semi-definite, and sparse when the input signals, \mathbf{x}_l , are sparse. Note that we have $\mathbf{X}_{k,l} \triangleq \mathbf{x}_k \mathbf{x}'_l \in \mathbb{C}^{N \times N}$. A naive approach to reconstruct the training signals is using CPRL with only the magnitude of the measurements $y_{i,k}$, i.e., by enforcing the constraints $|y_{i,k}|^2 = \mathbf{m}'_i \mathbf{X}_{k,k} \mathbf{m}_i$, $i = 1 \dots M$, $k = 1 \dots L$ as in (3). With this formulation, we have LM constraints to reconstruct L sparse rank-one matrices of size $N \times N$. However, we do not exploit the fact that the unknown phase shift factors are the same for *all* the L training signals.

To exploit these correlations, we define the cross measurements $g_{i,k,l}$ between the k^{th} and l^{th} measurements at the i^{th} sensor as

$$g_{i,k,l} \triangleq y_{i,k} y'_{i,l} \quad i = 1 \dots M \quad (6)$$

$$= e^{j\theta_i} \mathbf{m}'_i \mathbf{x}_k \mathbf{x}'_l \mathbf{m}_i e^{-j\theta_i} \quad k, l = 1 \dots L \quad (7)$$

$$= \mathbf{m}'_i \mathbf{X}_{k,l} \mathbf{m}_i. \quad (8)$$

Then, to reconstruct the original training signals, we propose to recover the joint matrix \mathbf{X} by solving

$$\begin{aligned} \mathbf{X}^* &= \arg \min_{\mathbf{Z}} \text{trace}(\mathbf{Z}) + \lambda \|\mathbf{Z}\|_1 & (9) \\ g_{i,k,l} &= \mathbf{m}'_i \mathbf{Z}_{k,l} \mathbf{m}_i & i = 1 \dots M \\ \mathbf{Z} &\succeq 0 & k, l = 1 \dots L. \end{aligned}$$

One can remark that the optimization problem in (3) and (9) are both Quadratic Basis Pursuit. Indeed, in both cases, we want to reconstruct a sparse rank-one matrix. However, the number of constraints is multiplied by L because we now exploit the correlation between the measurements. Even though the input signals are assumed to be sparse, the problem can easily be modified to handle the cases where the signal is sparse in a known domain, Ψ , such that $\|\mathbf{Z}\|_1$ is replaced with $\|\Psi \mathbf{Z} \Psi'\|_1$ in (9).

Finally, note that, as in (2) and (3), the final estimated signal \mathbf{x}^* (and therefore $\mathbf{x}^*_1, \dots, \mathbf{x}^*_L$) is defined up to a global phase since $\mathbf{X}^* = \mathbf{x}^* \mathbf{x}^{*\prime}$. The phases θ_i can be recovered given $y_{i,l}$ and \mathbf{x}^* .

2.3. Scalable Phase Calibration

One can observe that both the size of \mathbf{X} and the number of cross measurements grow with L^2 which can severely affect the scalability of the proposed algorithm. To address this problem, we propose instead to construct sub-matrices of \mathbf{X} :

$$\bar{\mathbf{X}}_l \triangleq \begin{bmatrix} \mathbf{X}_{l,l} & \mathbf{X}_{l,(l)L+1} \\ \mathbf{X}_{(l)L+1,l} & \mathbf{X}_{(l)L+1,(l)L+1} \end{bmatrix} \quad l = 1 \dots L \quad (10)$$

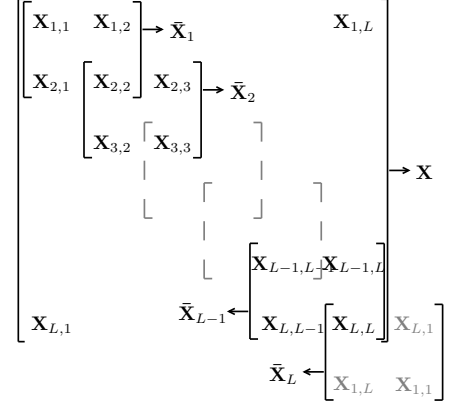


Fig. 1: The sub-matrices $\bar{\mathbf{X}}_1, \dots, \bar{\mathbf{X}}_L$ within the matrix \mathbf{X} .

where $(l)_L \triangleq l \pmod{L}$. A depiction of the sub-matrices with respect to \mathbf{X} can be seen in Figure 1. Similarly to \mathbf{X} , $\bar{\mathbf{X}}_l$ are also rank-one, positive semi-definite and sparse matrices. Therefore, an alternative to the optimization problem (9) is to construct only the sub-matrices by solving

$$\begin{aligned} \bar{\mathbf{X}}_1^*, \dots, \bar{\mathbf{X}}_L^* &= \\ \arg \min_{\bar{\mathbf{Z}}_1 \dots \bar{\mathbf{Z}}_L} & \sum_{k=1}^L \text{trace}(\bar{\mathbf{Z}}_k) + \lambda \|\bar{\mathbf{Z}}_k\|_1 & (11) \\ \text{subject to} & \bar{\mathbf{Z}}_1 \succeq 0, \dots, \bar{\mathbf{Z}}_L \succeq 0 \\ i = 1 \dots M & g_{i,l,l} = \mathbf{m}'_i \bar{\mathbf{Z}}_{l,l} \mathbf{m}_i \\ l = 1 \dots L & g_{i,(l)_L+1,l} = \mathbf{m}'_i \bar{\mathbf{Z}}_{(l)_L+1,l} \mathbf{m}_i \\ & g_{i,l,(l)_L+1} = \mathbf{m}'_i \bar{\mathbf{Z}}_{l,(l)_L+1} \mathbf{m}_i \\ & \bar{\mathbf{Z}}_l = \begin{bmatrix} \mathbf{Z}_{l,l} & \mathbf{Z}_{l,(l)_L+1} \\ \mathbf{Z}_{(l)_L+1,l} & \mathbf{Z}_{(l)_L+1,(l)_L+1} \end{bmatrix} \end{aligned}$$

As a result of this optimization, the estimated signals can be recovered considering

$$\bar{\mathbf{X}}_l^* = \begin{bmatrix} \mathbf{x}_l^* \\ \mathbf{x}_{(l)_L+1}^* \end{bmatrix} \begin{bmatrix} \mathbf{x}_l^{*\prime} & \mathbf{x}_{(l)_L+1}^{*\prime} \end{bmatrix} \quad l = 1 \dots L \quad (12)$$

Note that in order to recover the input signals up to a global phase shift, the phases of each input signal $\mathbf{x}_l, l = 1, \dots, L$ must be adjusted so that all signals are consistent with the sub-matrices $\bar{\mathbf{X}}_1, \dots, \bar{\mathbf{X}}_L$.

The optimization in (11) deals with L matrices of size $2N \times 2N$ instead of a single $LN \times LN$ matrix. The number of constraints are also reduced from ML^2 to $3ML$ for $L > 3$ with respect to (9) while retaining the joint recovery characteristic. A clear comparison can be seen in Table 1.

There are several remarks that can be made on the joint phase retrieval and calibration methods in (9) and (11):

1. The proposed methods can only recover the phase shifts at the sensors and the input signals up to a global phase shift factor. However the global phase shift is common

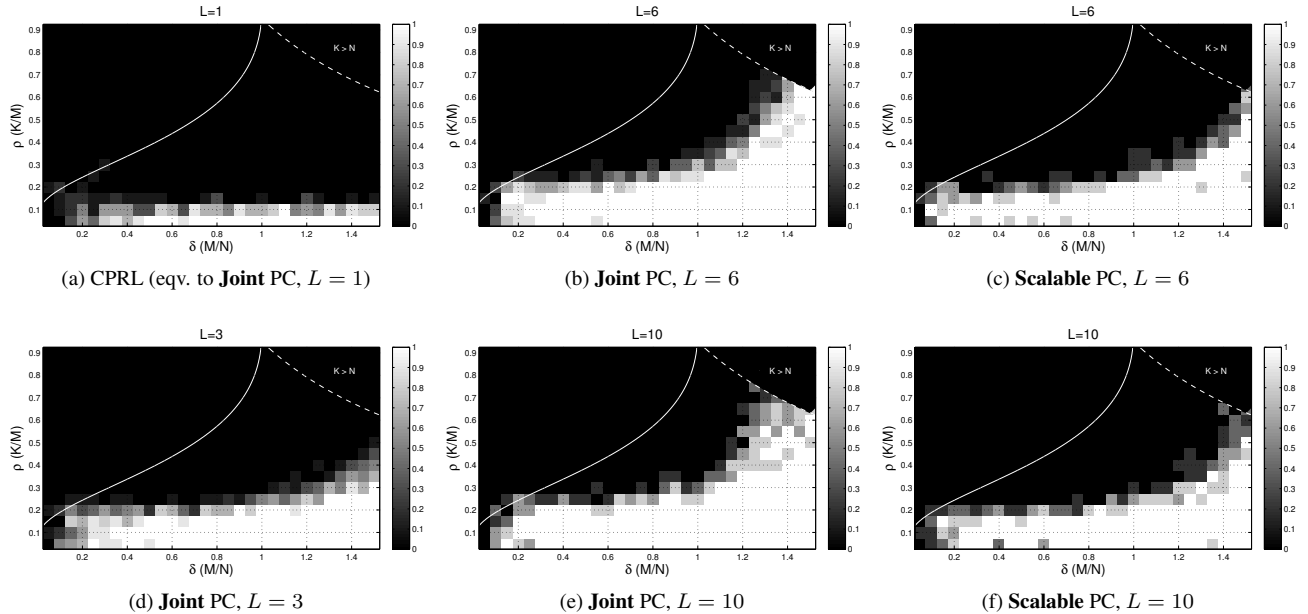


Fig. 2: The probability of perfect recovery for $N = 100$ with respect to $\delta \triangleq M/N$ and $\rho \triangleq K/M$. The solid line indicates the Donoho-Tanner phase transition curve for fully calibrated compressed sensing recovery [10]. The dashed line indicates the boundary to the region where $K > N$.

for all θ_i and \mathbf{x}_l unlike the case when \mathbf{x}_l are recovered by CPRL individually using (3).

- Even though the optimization method utilized in (9) is similar to (3), the joint measurements, $g_{i,k,l}$, are complex valued and the phase information in $y_{i,l}$ are utilized in the recovery. While this improves the performance, the proposed method cannot be applied to cases where only the magnitude of the measurements are directly observed or phase information is completely lost.
- The matrix \mathbf{X} has a much larger size than the matrices in Phaselift or CPRL which results in higher computational complexity and memory requirement. This issue can partly be overcome with algorithms exploiting the sparsity and low rank nature of the matrix during optimization to reduce the memory requirements as well as complexity as discussed in [3]. The matrices $\bar{\mathbf{X}}_1, \dots, \bar{\mathbf{X}}_L$ have comparable size to the CPRL or Phaselift algorithm.

3. EXPERIMENTAL RESULTS

In order to test the performance of the proposed algorithm, phase transition curves as in the compressed sensing recovery are plotted for a signal size $N = 100$ with the measurement vectors, \mathbf{m}_i , and all the non zero entries in the input signals, \mathbf{x}_l , randomly generated from an i.i.d. normal distribution.

Table 1: The comparison between CPRL and the proposed methods for the reconstruction of L input signals of size N

	# of Constraints	Size of Unknown(s)	Final Output
CPRL	ML	L matrices of $N \times N$	L signals with L arbitrary phase shifts
Joint PC	ML^2	$NL \times NL$ matrix	L signals with single global phase shift
Scalable PC	$3ML$	L matrices of $2N \times 2N$	L signals with single global phase shift

The signals (and the phase shift parameters) are recovered for the number of inputs $L = 1, 3, 6, 10$ with the proposed optimization in (9) using an ADMM [11] based algorithm. The fast optimization method proposed in (11) is also performed for number of inputs, $L = 6, 10$, for comparison. The perfect reconstruction criteria is selected as $\sigma(\mathbf{x}, \mathbf{x}^*) > 0.9$, where the absolute correlation factor $\sigma(\cdot, \cdot)$ is defined as

$$\sigma(\mathbf{x}_1, \mathbf{x}_2) \triangleq \frac{|\mathbf{x}'_1 \mathbf{x}_2|}{\|\mathbf{x}_1\|_2 \|\mathbf{x}_2\|_2} \quad (13)$$

so that the global phase difference between the source and recovered signals is ignored.

It is observed in the simulations that the optimization parameter λ is quite sensitive and must be selected carefully. In our experiments we found that a good choice of lambda varies with $\rho = K/M$, and we used the experimentally found values

Table 2: Variation of λ with respect to $\rho = K/M$ in the experiments. The λ at remaining ρ values are linearly interpolated from nearest points in the table.

ρ	0.05	0.1	0.2	0.4	0.9
λ	2.5	0.4	0.1	0.03	0.01

shown in Table 2. A rule of thumb for selecting λ can be inferred from the set of values in Table 2 such that $\lambda \approx 0.04 \frac{1}{\rho^2}$. It should be noted that these values are optimized for higher speed of convergence whereas a much broader range of λ values permit recovery. Therefore a viable λ can be chosen without knowing K or ρ .

The probability of recovery of each of the proposed methods with respect to $\delta \triangleq M/N$ and $\rho \triangleq K/M$ are shown in Figure 2 for a varying number of input signals, L . The results provided for $L = 1$ represents the performance of the individual recovery of the input signals which is equivalent to the CPRL method [6]. It can be observed that the proposed joint recovery methods provide significantly better performance than individually recovering the signals, even when there are only few input signals. The performance keeps improving with increasing L , although the improvement gets less noticeable as L gets larger.

The performance of the scalable optimization can be seen in Figures 2(c) and 2(f). Although there is a slight performance degradation with respect to the joint optimization, the scalable optimization can still be preferable when L is too large and computational or memory requirements for joint optimization are too high. In our simulations, we observed that the scalable optimization took less than half the time and consumed around half the memory than the joint optimization for $L = 10$, however it should be noted that the convergence is observed to be slower in terms of number of iterations due to larger number of constraints in (11).

4. CONCLUSIONS AND FUTURE WORK

In this paper we have investigated sparse signal recovery in a compressive measurement system in which each sensor introduces an unknown but fixed phase shift to the measurements. Extending the phase retrieval methods studied in the literature, we have shown that it is possible to significantly improve the signal recovery performance when multiple unknown inputs are measured through the system. Our approach makes use of the correlation among the multiple measurements by each sensor to jointly recover the unknown sparse inputs and the phase shifts introduced by each sensor up to a global phase factor common to all unknowns. A second optimization approach with significantly better scalability at the expense of slightly reduced performance is also introduced. It has computational complexity and memory requirements that are linear in the number of sparse inputs which as a result is prefer-

able for large number of input signals.

As a future work, we first plan to investigate the limits of the performance of the presented algorithms as the number of input signals gets larger and in the presence of noise. Secondly, we plan on developing a more general approach to perform complex valued gain calibration combining the methods introduced in this paper with earlier methods studied for gain magnitude calibration. Lastly, our aim is to investigate non-convex approaches for faster and memory efficient recovery.

5. REFERENCES

- [1] David L. Donoho, "Compressed Sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289 – 1306, 2006.
- [2] Cagdas Bilen, Gilles Puy, Rémi Gribonval, and Laurent Daudet, "Blind Sensor Calibration in Sparse Recovery Using Convex Optimization," in *10th International Conference on Sampling Theory and Applications (SAMPTA)*, 2013.
- [3] Emmanuel J. Candes, Thomas Strohmer, and Vladislav Voroninski, "PhaseLift: Exact and stable signal recovery from magnitude measurements via convex programming," *Communications on Pure and Applied Mathematics (to appear)*, 2011.
- [4] Radu Balan, Pete Casazza, and Dan Edidin, "On signal reconstruction without phase," *Applied and Computational Harmonic Analysis*, vol. 20, no. 3, pp. 345–356, May 2006.
- [5] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of the phase from image and diffraction plane pictures," *Optik*, vol. 35, no. 2, pp. 237–246, 1972.
- [6] Henrik Ohlsson, Allen Y Yang, Roy Dong, and Shankar S. Sastry, "Compressive phase retrieval from squared output measurements via semidefinite programming," *arXiv preprint arXiv:1111.6323*, 2011.
- [7] Henrik Ohlsson, Allen Yang, Roy Dong, and Shankar Sastry, "Quadratic Basis Pursuit," in *Signal Processing with Adaptive Sparse Structured Representations (SPARS) Workshop*, Lausanne, Switzerland, 2013.
- [8] J.T. Parker, V. Cevher, and P. Schniter, "Compressive sensing under matrix uncertainties: An approximate message passing approach," in *Signals, Systems and Computers (ASILOMAR), 2011 Conference Record of the Forty Fifth Asilomar Conference on*, nov. 2011, pp. 804–808.
- [9] Philip Schniter and Sundeep Rangan, "Compressive Phase Retrieval via Generalized Approximate Message Passing," in *Proceedings of Allerton Conference on Communication, Control, and Computing*, 2012.
- [10] David L. Donoho and Jared Tanner, "Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing.," *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, vol. 367, no. 1906, pp. 4273–93, Nov. 2009.
- [11] Stephen Boyd, Neal Parikh, Eric Chu, B Peleato, and Jonathan Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.