Convex Optimization Approaches for Blind Sensor Calibration using Sparsity

Cagdas Bilen, Gilles Puy, Rémi Gribonval, Laurent Daudet

Convex Optimization Approaches for Blind Sensor Calibration using Sparsity

Çağdaş Bilen*, Gilles Puy†, Rémi Gribonval* and Laurent Daudet‡
* INRIA, Centre Inria Rennes - Bretagne Atlantique, 35042 Rennes Cedex, France.
† Inst. of Electrical Eng., 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 investigate a compressive sensing framework in which the sensors introduce a distortion to the measurements in the form of unknown gains. We focus on blind calibration, using measurements on multiple unknown (but sparse) signals and formulate the joint recovery of the gains and the sparse signals as a convex optimization problem. The first proposed approach is an extension to the basis pursuit optimization which can estimate the unknown gains along with the unknown sparse signals. Demonstrating that this approach is successful for sufficient number of input signals except in cases where the phase shifts among the unknown gains varies significantly, a second approach is proposed that makes use of quadratic basis pursuit optimization to calibrate for constant amplitude gains with maximum variance in the phases. An alternative form of this approach is also formulated to reduce the complexity and memory requirements and provide scalability with respect to the number of input signals. Finally a third approach is formulated which combines the first two approaches for calibration of systems with any variation in the gains. The performance of the proposed algorithms are investigated extensively through numerical simulations, which demonstrate that simultaneous signal recovery and calibration is possible when sufficiently many (unknown, but sparse) calibrating signals are provided.**

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

I. Introduction

Compressed sensing is a theoretical and numerical framework to sample sparse signals at lower rates than required by the Nyquist-Shannon theorem [1]. More precisely, a K-sparse source vector $\mathbf{x} \in \mathbb{C}^N$ is sampled by a number M of linear measurements

$$y_i = \mathbf{m}_i' \mathbf{x}, \qquad i = 1, \dots, M \tag{1}$$

where $\mathbf{m}_1, \dots, \mathbf{m}_M \in \mathbb{C}^N$ 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.,

$$\hat{\mathbf{x}}_{\ell_1} = \arg\min_{\mathbf{z}} \|\mathbf{z}\|_1 \tag{2}$$

$$\text{subject to} \quad y_i = \mathbf{m}_i' \mathbf{z}, \quad i = 1, \dots, M \tag{3}$$

where $\|\cdot\|_1$ denotes the ℓ_1-norm, which favors the selection of sparse signals among the ones satisfying the measurement constraints. It has been showed that the number of measurements needed for accurate recovery, such that $\hat{\mathbf{x}}_{\ell_1} = \mathbf{x}$, scales only linearly with K [1]. Note that the above minimization problem can easily be modified to handle the presence of additive noise on the measurements.

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.

A. Calibration in Inverse Problems

In some practical situations, it is not possible to perfectly know the measurement vectors $\mathbf{m}_1, \dots, \mathbf{m}_M$. In many applications dealing with distributed sensors or radars, the location or intrinsic parameters of the sensors are not exactly known, which in turn results in unknown phase shifts and/or gains at each sensor [2], [3]. Similarly, applications with microphone arrays are shown to require calibration of each microphone to account for the unknown gain and phase shifts introduced [4]. Recently different blind calibration scenarios for the measurement matrix in compressive sensing has been considered. Perturbations in a parameterized measurement matrix are estimated along with the signal in [5]. Unlike additive perturbations in the measurement matrix, compressive sensing reconstruction is not robust to multiplicative perturbation [6], [7]. Generalized approximate message passing (GAMP) algorithm is used to estimate unknown scaling for the input signal [8] or the unknown gains introduced by the sensors [9] along with the unknown sparse signal(s). We propose in this paper various convex optimization methods to automatically calibrate the measurement system while estimating the unknown sparse input signals in the latter scenario.

B. Phase Retrieval

Some of the methods proposed here are related to techniques developed to solve the phase retrieval problem, which occurs in imaging techniques such as optical interferometric imaging for astronomy. In this problem, one has only access to the magnitude of the measurements $z_i = |y_i|^2 = \mathbf{m}_i' \mathbf{x}\mathbf{x}' \mathbf{m}_i$, $i = 1, \dots, M$, where $\mathbf{m}_1, \dots, \mathbf{m}_M$ are vectors of the Fourier basis. Reconstructing the original signal from such magnitude measurements is a phase retrieval problem and seems more challenging than solving problem (1). This problem is interesting as it is a specific case of the calibration problem with unknown phase shifts at each sensor resulting in absolute value of the measurements. Nevertheless, Candès *et al.* have recently showed [10], [11] that $\mathbf{x}$ could be recovered exactly by solving a convex optimization problem with a number of measurements, $M > N$, essentially proportional to N.

Instead of directly looking for a signal vector $\mathbf{x}$, the method relies on finding a positive semi-definite matrix $\mathbf{X} \triangleq \mathbf{x}\mathbf{x}'$ of rank-one such that $|y_i|^2 = \mathbf{m}_i'\mathbf{X}\mathbf{m}_i$ for all $i = 1, \ldots, M$. The measurement constraint becomes linear and Candès *et al.* propose to solve the following convex optimization called the Phaselift (PL) to recover $\mathbf{X}$:

$$\begin{aligned} \hat{\mathbf{X}}_{\text{PL}} = \arg\min_{\mathbf{Z}} \quad & \text{Tr}(\mathbf{Z}) \\ \text{subject to} \quad & \mathbf{Z} \succcurlyeq 0 \\ & |y_i|^2 = \mathbf{m}_i'\mathbf{Z}\mathbf{m}_i, \quad i = 1, \ldots, M \end{aligned} \tag{4}$$

The trace norm $\text{Tr}(\cdot)$ favors the selection of low rank matrices among all the ones satisfying the constraints. Let us acknowledge that this phase retrieval problem was also previously studied theoretically in, e.g., [12], [13], but a larger number of measurements is needed for reconstruction of the original signal with the technique therein ($M \propto N^2$ instead of $M \propto N$). Note also that several simple iterative algorithms such as the one described in [14] 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.* [15], [16]. This new approach is called Compressive Phase Retrieval via Lifting (CPRL) or Quadratic Basis Pursuit [17] and consists in solving the problem (4) with the addition of a cost term that penalizes non-sparse matrices. This extra term allows them to reduce the number of magnitude measurements needed to accurately recover the sparse signals. The convex optimization becomes

$$\begin{aligned} \hat{\mathbf{X}}_{\text{CPRL}} = \arg\min_{\mathbf{Z}} \quad & \text{Tr}(\mathbf{Z}) + \lambda\|\mathbf{Z}\|_1 \\ \text{subject to} \quad & \mathbf{Z} \succcurlyeq 0 \\ & |y_i|^2 = \mathbf{m}_i'\mathbf{Z}\mathbf{m}_i, \quad i = 1, \ldots, M, \end{aligned} \tag{5}$$

where $\lambda > 0$. The authors also provide bounds for guaranteed recovery of this method using a generalization of the restricted isometry property. Note that conditions on the number of measurements for accurate reconstruction of sparse signals by this approach when the measurements are drawn randomly from the normal distribution are also available in [18].

C. Our Contribution

In this paper, we investigate several convex optimization approaches for the joint recovery of unknown complex valued gains introduced by the sensors along with unknown sparse signals in the compressed sensing system described in Section II. In Section III, we introduce a convex optimization approach that can jointly recover the unknown gains and sparse input signals when the sign or phase variation among the gains is small. An alternative approach whose performance is independent of the variation in phases of the gains is introduced in Section IV. This approach is an adaptation of the Quadratic Basis Pursuit to the compressive calibration problem. A combination of these two methods is then introduced in Section VI to handle the case where both the amplitude and the phase decalibration are large. Conclusions and discussions on future work is presented in Section VII.

II. Problem Definition

Suppose that the measurement system in (1) is perturbed by complex gains, $d_i e^{j\theta_i}$, at each sensor i and that multiple sparse input signals, $\mathbf{x}_\ell \in \mathbb{C}^N, \quad \ell = 1 \ldots L$, are measured through the system such that

$$y_{i,\ell} = d_i e^{j\theta_i}\mathbf{m}_i'\mathbf{x}_\ell \qquad i = 1, \ldots, M,\ \theta_i \in [0, 2\pi),\ d_i \in \mathbb{R}^+ \tag{6}$$

For a real valued system, a similar problem can be considered with signed real gains for which the phase term $e^{j\theta_i}$ is replaced by ± 1. We will treat the unknown signs in the real valued systems as unknown phases with $\theta_i \in \{0, \pi\}$, and use the same terminology for both the real and complex valued systems for the rest of this paper. In a compressed sensing scenario the signal is estimated from incomplete measurements ($M < N$), however in the calibration problem considering a larger number of measurements ($M > N$) is also of interest since there is a larger number of unknowns therefore we consider both cases. We focus only on the noiseless case for the sake of simplicity.

It has been shown that ignoring the unknown gains during recovery, i.e. treating the distortion as noise, can significantly reduce recovery performance especially when there is ambiguity in the phase [7]. Therefore it is essential to employ a reconstruction approach that deals with unknown gains rather than ignoring them.

In a traditional recovery strategy, one can enforce the sparsity of the input signals while enforcing the measurement constraints in (6). However, when dealing with unknown gains, the measurement constraints are non-linear with respect to the unknowns d_i and $\mathbf{x}_\ell$. This non-linearity can be tackled using an alternate minimization strategy where one iteratively estimates $\mathbf{x}$ while keeping d_i fixed and vice-versa [2]. However, the convergence of this alternating optimization to the global minimum is not guaranteed.

In this paper, we formulate the problem of joint signal recovery and calibration from the measurements $y_{i,\ell}$ in (6) as a convex optimization problem. We shall consider three main scenarios on the sensor characteristics:

1) **Amplitude Calibration**: small or no variation in phase shifts among the sensors
2) **Phase Calibration**: high variation in unknown phase shifts among the sensors but with known sensor amplification
3) **Complete Calibration**: high variation in unknown phase shifts among the sensors and unknown sensor amplification

each of which will be approached and formulated differently.

III. Amplitude Calibration

In this section, we first consider the case where the phase decalibration at each sensor is null (or, equivalently, known), i.e., $\theta_i = 0$ and $d_i \geq 0$. As proposed in [19], it is then possible to overcome the non-linearity in the measurement equation (6) with respect to d_i and $\mathbf{x}_\ell$ by re-parameterizing it in a bi-linear

fashion such that

$$y_{i,\ell}\tau_i = \mathbf{m}_i'\mathbf{x}_\ell \qquad i = 1,\dots,M\ ,\ \ell = 1,\dots,L \qquad (7)$$
$$\tau_i \triangleq \frac{1}{d_i}$$

assuming that $d_i \neq 0\ \forall i$ without loss of generalization. Consequently, one can attempt to recover the sparse signals and the gains with the convex optimization approach

***A-Cal*:**

$$\begin{matrix}\hat{\mathbf{x}}_1,\dots,\hat{\mathbf{x}}_L \\ \hat{\tau}_1,\dots,\hat{\tau}_M\end{matrix} = \underset{\substack{\mathbf{z}_1,\dots,\mathbf{z}_L \\ t_1,\dots,t_M}}{\arg\min} \sum_{n=1}^{L}\|\mathbf{z}_n\|_1 \qquad (8)$$
$$\text{subject to} \quad y_{i,\ell}t_i = \mathbf{m}_i'\mathbf{z}_\ell \qquad \begin{matrix}\ell = 1,\dots,L \\ i = 1,\dots,M\end{matrix}$$
$$\sum_{n=1}^{M} t_n = c$$

for an arbitrary constant $c > 0$. The actual gains are recovered after the optimization using $\hat{d}_i = \frac{1}{\hat{\tau}_i}$.

The optimization ***A-Cal*** retrieves the smallest ℓ_1 norm solution for the $NL + M$ unknowns satisfying the $ML + 1$ constraints. Assuming the measurement equations are consistent and $\mathbf{m}_1,\dots,\mathbf{m}_M$ span $\mathbb{C}^N$, the given constraints can be satisfied with a unique solution provided that $M \geq \frac{NL-1}{L-1} \cong N(1 + \frac{1}{L-1})$. If this condition is satisfied, a solution can be reached more robustly in the presence of noise with a least squares estimation

***LS-Cal*:**

$$\begin{matrix}\hat{\mathbf{x}}_1,\dots,\hat{\mathbf{x}}_L \\ \hat{\tau}_1,\dots,\hat{\tau}_M\end{matrix} = \underset{\substack{\mathbf{z}_1,\dots,\mathbf{z}_L \\ t_1,\dots,t_M}}{\arg\min} \sum_{\ell=1}^{L}\sum_{i=1}^{M}\|y_{i,\ell}t_i - \mathbf{m}_i'\mathbf{z}_\ell\|_2^2 \qquad (9)$$
$$\text{subject to} \quad \sum_{n=1}^{M} t_n = c$$

for which a closed form solution can be derived.

We can make several observations regarding the optimization ***A-Cal***:

1) The constraint $\sum_n t_n = c$ ensures that the trivial solution ($\tau_i = 0,\ \mathbf{x}_\ell = 0,\ \forall i,\ell$) is excluded from the solution set.
2) The constraint $\sum_n t_n = c$ also excludes the solutions where the sum of the gains are zero. When dealing with gains of varying phase shifts among the sensors, this may result in excluding the actual solution in rare cases where the sought out gains actually sum up to zero.
3) The measurement constraints are satisfied up to a global scale factor and phase shift, therefore the constant c can be set arbitrarily.
4) The successful recovery of the gains and the signals require availability of more than one input signal ($L > 1$). Although this may seem like a restriction, acquiring data with the same sampling system from multiple sources is often straightforward in many application fields.
5) The bound, $M \geq \frac{NL-1}{L-1}$, for reaching a closed form solution with ***LS-Cal*** demonstrates another benefit of *joint* optimization, since the bound is finite only for $L > 1$.

The optimization approach ***A-Cal*** has been investigated for $\theta_i = 0$ and various experimental results are presented [19]. However it is also possible to use the same approach for systems with varying phase shifts or sign changes among the gains. Therefore we shall consider the same optimization approach for the calibration of systems with varying characteristics in the distribution of the phases shifts in the gains, for which the experimental results are presented in the following section.

A. Experimental Results for Amplitude Calibration

1) Data Generation: In order to test the performance of the amplitude calibration approach ***A-Cal***, 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 K non-zero entries in the input signals, $\mathbf{x}_\ell$, randomly generated from an i.i.d. normal distribution. The positions of the non-zero coefficients of the input signals, $\mathbf{x}_\ell$, are chosen uniformly at random in $\{1,\dots,N\}$. The magnitude of the gains were generated using $\log d_i \sim \mathcal{N}(0,\sigma^2)$, where σ is the parameter governing the range of decalibration. The phase of the gains are chosen uniformly at random from the range $[0\ 2\pi p_c)$ where $p_c \in \{0, 0.33, 0.66, 1\}$. Note that the parameter p_c determine the scale of ambiguity in the phases where maximum possible ambiguity is observed for $p_c = 1$.

The signals (and the gains) are recovered for different ranges of amplitude decalibration ($\sigma = 0.1,\ 0.3,\ 1$) with a sufficiently high number of input signals ($L = 5,\ 10,\ 20$ respectively). For each parameter set, 10 randomly generated set of signals and gains are recovered by the proposed optimization ***A-Cal***, which is performed using an ADMM [21] algorithm.

2) Performance Measure: The perfect reconstruction criteria is selected as $\frac{1}{L}\sum_\ell \mu(\mathbf{x}_\ell, \hat{\mathbf{x}}_\ell) > 0.9999$, where the absolute correlation factor $\mu(\cdot,\cdot)$ is defined as

$$\mu(\mathbf{x}_1, \mathbf{x}_2) \triangleq \frac{|\mathbf{x}_1'\mathbf{x}_2|}{\|\mathbf{x}_1\|_2\|\mathbf{x}_2\|_2} \qquad (10)$$

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

3) Results: The probability of recovery (computed through 10 independent simulations for each set of parameters) of the proposed method with respect to $\delta \triangleq M/N$ and $\rho \triangleq K/M$ are shown in Figure 1. The first thing to notice from the results is that the performance in the absence of phase ambiguity ($p_c = 0$) is consistent with the results presented in [19] as expected. The effect of increasing phase ambiguity can be observed in the results as p_c increases. Although the performance is acceptable for p_c as high as 0.66, there is a significant degradation when dealing very high phase ambiguity such that signal recovery is impossible regardless of the sparsity, unless the measurement system is overcomplete ($M > N$). This phenomena can best be observed in the last row of results in Figure 1, where the number of input signals is very large ($L = 50$) with respect to the variance in the

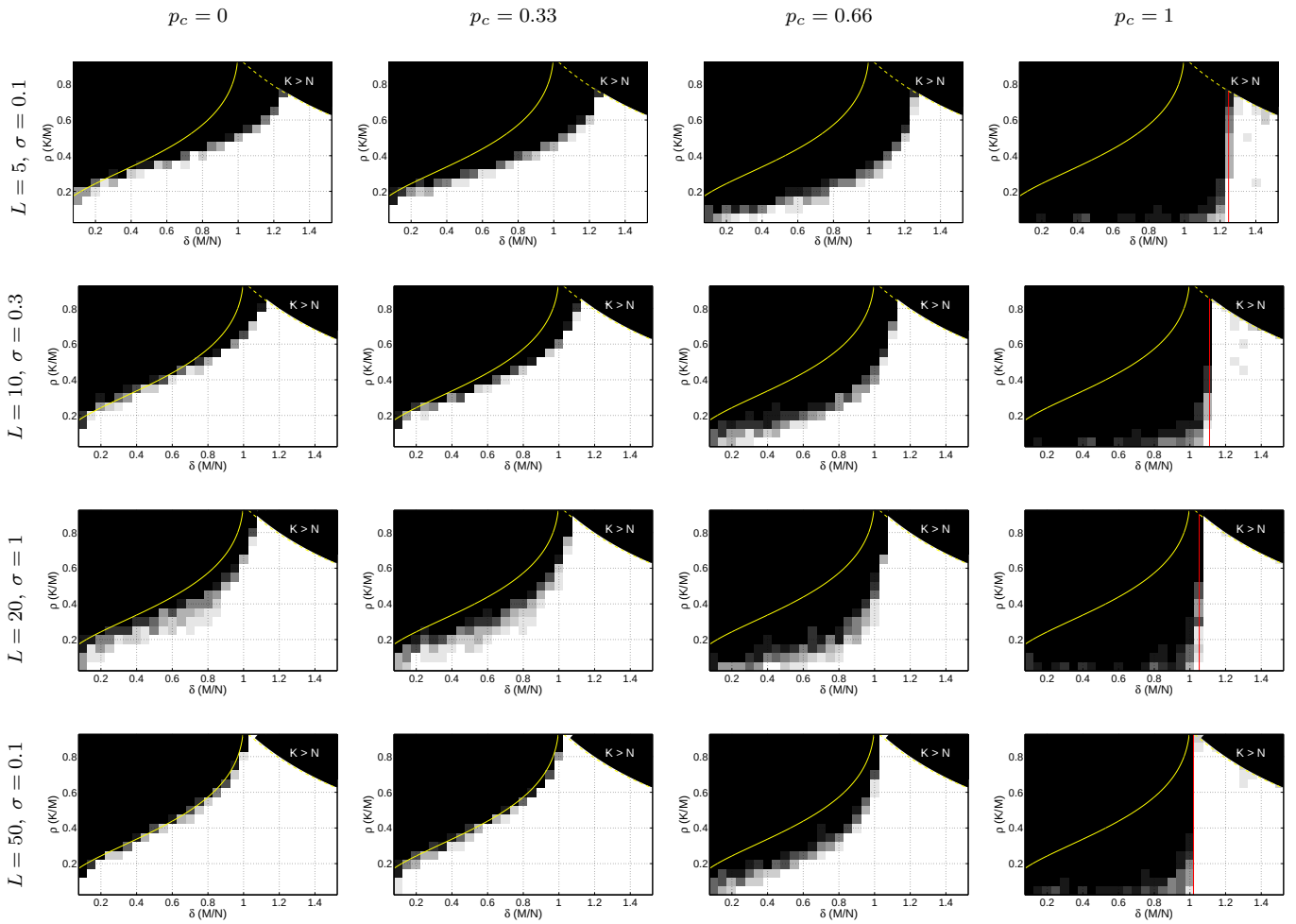


Fig. 1: (***A-Cal***) The probability of perfect recovery (colors black to white indicate probabilities 0 to 1) in the complex valued system for $N = 100$ with respect to $\delta \triangleq M/N$ and $\rho \triangleq K/M$. The solid yellow line indicates the Donoho-Tanner phase transition curve for fully calibrated compressed sensing recovery [20]. The dashed yellow line indicates the boundary to the unfeasible region where $K > N$. Each row of figures display the change in recovery performance with increasing phase ambiguity from left to right for a fixed set of L and σ. The last row shows the performance limit for very high L. The red line in the rightmost column indicates the bound $\delta = \frac{NL-1}{NL-N}$ for unique recovery without optimization.

gain magnitudes ($\sigma = 0.1$). Very similar performance for real valued systems with varying sign ambiguity in the gains is observed in experiments for which the results are not shown here.

The minimum number, L, of input signals necessary to obtain a successful recovery of all 10 of the example input signals depends on the range of decalibration, σ, as well as the phase variability, p_c, as can be observed on Figure 2. For small variations on the phase ($p_c = 0, 0.2, 0.4$) only a few input signals are sufficient unless there is very high variation in the amplitude. However as phase variation increases ($p_c > 0.6$), very large number of input signals are required regardless of variation in the amplitude.

It is observed that the minimum number of input signals required for recovery remains almost constant for increasing size of the signal ($L = 4, 3, 3$ for $N = 100, 1000, 5000$ respectively), when all the other parameters are kept constant ($p_c = 0.5$, $\sigma = 0.3$, $\rho = 0.2$, $\delta = 0.8$).

B. Discussion

The diminishing performance in Figure 1 as the phase variation, p_c, increases can be attributed to the significant increase in the contamination of the information in the measurements as the phase ambiguity increases. Therefore recovery becomes possible only when there are sufficiently many measurements. For the maximally ambiguous case ($p_c = 1$) the results show that the proposed optimization is not needed since it fails unless $M \geq \frac{NL-1}{L-1}$, in which case a more convenient closed form solution can be derived. Even though this is a drawback of the presented approach, it should be noted that in many practical systems the phase ambiguity is restricted to a small range. Such scenarios represent typical use cases for this algorithm.

We derive in the next section an alternative approach, formulated via Quadratic Basis Pursuit, to deal with cases where the phase variation in the unknown gains is significant.

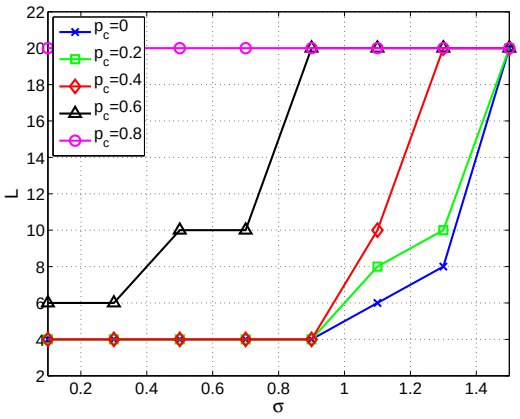


Fig. 2: (***A-Cal***) Minimum number of training signals, L, necessary for successful recovery, as a function of the range of decalibration σ, for various levels of phase ambiguity p_c, when $\rho = 0.2$, $\delta = 0.8$ and $N = 100$. Successful recovery is defined as recovery of all 10 randomly generated set of signals and gains for the fixed parameters.

IV. Phase Calibration

As an alternative to direct recovery described in Section III, it is possible to eliminate the phase ambiguity in the measurements by creating quadratic measurements similar to the approach by Candes *et al.* [11] and Ohlsson *et al.* [15], [17] used for phase retrieval problems. Since the unknown phase shifts are removed from the new optimization problem, it is possible to have signal recovery and calibration even when there is a high variance in phase shifts among the sensors. In this section, we consider the scenario of known gain amplitudes, hence $d_i\mathbf{m}_i$ is simply replaced with $\mathbf{m}_i$ for the rest of the discussions.

Let us define the cross measurements, $g_{i,k,\ell}$ as

$$g_{i,k,\ell} \triangleq y_{i,k}y'_{i,\ell} \qquad i = 1,\dots,M,\ k,\ell = 1,\dots,L \tag{11}$$

$$= e^{j\theta_i}\mathbf{m}'_i\mathbf{x}_k\mathbf{x}'_\ell\mathbf{m}_i e^{-j\theta_i} \tag{12}$$

$$= \mathbf{m}'_i\mathbf{X}_{k,\ell}\mathbf{m}_i \qquad \mathbf{X}_{k,\ell} \triangleq \mathbf{x}_k\mathbf{x}'_\ell \in \mathbb{C}^{N\times N} \tag{13}$$

We can also define the joint signal matrix $\mathbf{X} \in \mathbb{C}^{LN\times LN}$

$$\mathbf{X} \triangleq \underbrace{\begin{bmatrix}\mathbf{x}_1\\ \vdots \\ \mathbf{x}_L\end{bmatrix}}_{\mathbf{x}} \underbrace{\begin{bmatrix}\mathbf{x}'_1 & \cdots & \mathbf{x}'_L\end{bmatrix}}_{\mathbf{x}'} = \begin{bmatrix}\mathbf{X}_{1,1} & \cdots & \mathbf{X}_{1,L}\\ \vdots & \ddots & \vdots \\ \mathbf{X}_{L,1} & \cdots & \mathbf{X}_{L,L}\end{bmatrix} \tag{14}$$

which is rank-one, hermitian, positive semi-definite and sparse when the input signals, $\mathbf{x}_\ell$, are sparse. Therefore we propose to recover the joint matrix $\mathbf{X}$ with the semi-definite program

***P-Cal*:**

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{Z}} \quad f_\lambda(\mathbf{Z}) \tag{15}$$

$$\text{subject to} \quad \mathbf{Z} \succcurlyeq 0$$

$$g_{i,k,\ell} = \mathbf{m}'_i\mathbf{Z}_{k,\ell}\mathbf{m}_i \qquad i = 1,\dots,M,\ k,\ell = 1,\dots,L$$

where

$$f_\lambda(\mathbf{Z}) \triangleq \mathrm{Tr}(\mathbf{Z}) + \lambda\|\mathbf{Z}\|_1 \tag{16}$$

One can remark that the optimization problem in (5) and (15)-(16) 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, $\mathbf{\Psi}$, such that $\|\mathbf{Z}\|_1$ is replaced with $\|\mathbf{\Psi}\mathbf{Z}\mathbf{\Psi}'\|_1$ in (16).

As in (4) and (5), the estimated signal, $\hat{\mathbf{x}}$ (and therefore $\hat{\mathbf{x}}_1,\dots.\hat{\mathbf{x}}_L$) is defined up to a global phase since $\hat{\mathbf{X}} = \hat{\mathbf{x}}\hat{\mathbf{x}}'$. The phases θ_i can be recovered given $y_{i,\ell}$ and $\hat{\mathbf{x}}$.

Unlike the standard 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 the use of 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 [11], [15], [22]. 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 each sensor i are *correlated* through the *same* phase shift factor, our approach exploits these correlations and phase information to reach better recovery performance than one could obtain though individual recovery such as with CPRL.

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.

A. Determining the Regularization Parameter λ

The choice of the parameter λ is essential for the success of the proposed Phase Calibration methods. For real valued systems, the bounds on M and λ for the CPRL recovery are investigated in [18] where sufficient conditions for perfect recovery is given as (assuming $\|\mathbf{x}\|_2 = 1$ without loss of generality) $\lambda > \sqrt{K}\|\mathbf{x}\|_1 + 1$, $\lambda < N^2/4$ and $M > C_0\lambda^2 \log N$ where C_0 is a constant. However our initial experiments has shown that these bounds are far from being tight, at least for signals of small sizes ($N = 100$).

As a separate approach to that of finding theoretical bounds on λ (and/or M), it is also possible to numerically determine the range of values for λ for which perfect recovery is

Algorithm 1 ***P-Cal-λ***: Determine if perfect recovery of $\mathbf{x}$ is possible and find upper and lower bounds on λ

Set recovery $\leftarrow$ **false**, $\lambda_{\text{low}} \leftarrow 0$, $\lambda_{\text{up}} \leftarrow \infty$
Perform optimization to find $\hat{\mathbf{D}}_0$ given $\mathbf{x}, \mathbf{m}_1, \ldots, \mathbf{m}_M$
if $\mathrm{G}(\hat{\mathbf{D}}_0) \leq 0$ **then**
  return (*recovery*, λ_{low}, λ_{up})
end if
Perform optimization to find $\hat{\mathbf{D}}_{-1}$ given $\mathbf{x}, \mathbf{m}_1, \ldots, \mathbf{m}_M$
if $\mathrm{G}(\hat{\mathbf{D}}_{-1}) \leq 0$ **then**
  return (*recovery*, λ_{low}, λ_{up})
else
  $\lambda_{\text{low}} \leftarrow \frac{1}{\mathrm{G}(\hat{\mathbf{D}}_{-1})}$
end if
Perform optimization to find $\hat{\mathbf{D}}_1$ given $\mathbf{x}, \mathbf{m}_1, \ldots, \mathbf{m}_M$
if $\mathrm{G}(\hat{\mathbf{D}}_1) < 0$ **then**
  $\lambda_{\text{up}} \leftarrow -\frac{1}{\mathrm{G}(\hat{\mathbf{D}}_1)}$
end if
if $\lambda_{\text{low}} < \lambda_{\text{up}}$ **then**
  recovery $\leftarrow$ **true**
end if
return (*recovery*, λ_{low}, λ_{up})

guaranteed *given* $\mathbf{x}$. An algorithm to determine if the perfect recovery is possible as well as the upper and lower bounds on the parameter λ for given $\mathbf{x}$ and $[\mathbf{m}_1, \ldots, \mathbf{m}_M]$ is shown in Algorithm 1 (***P-Cal-λ***). The term $\hat{\mathbf{D}}_p$ in Algorithm 1 represents the result of the optimization

$$\begin{aligned} \hat{\mathbf{D}}_p \triangleq \arg\min_{\mathbf{Z}} \quad & \mathrm{G}(\mathbf{Z}) && (17)\\ \text{subject to} \quad & \mathbf{m}_i' \mathbf{Z}_{k,\ell} \mathbf{m}_i = 0, && i = 1, \ldots, M \\ & \mathrm{Tr}(\mathbf{Z}) = p, && k, \ell = 1, \ldots, L \\ & \mathbf{Z} = \mathbf{E} \begin{bmatrix} a & \mathbf{b}' \\ \mathbf{b} & \mathbf{C} \end{bmatrix} \mathbf{E}', && a \in \mathbb{R},\ \mathbf{b} \in \mathbb{C}^{LN-1} \\ & \mathbf{C} \succcurlyeq 0, && \mathbf{C} \in \mathbb{C}^{LN-1 \times LN-1} \end{aligned}$$

where $\mathbf{E}$ is defined with the eigen-decomposition of $\mathbf{X} \triangleq \mathbf{x}\mathbf{x}'$ such that $\mathbf{X} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}'$. The function $\mathrm{G}(.)$ is defined as

$$\mathrm{G}(\mathbf{Z}) \triangleq \|\mathbf{Z}_{\Omega_X^\perp}\|_1 + \mathrm{Real}\{\langle \mathrm{sign}(\mathbf{X}), \mathbf{Z}_{\Omega_X} \rangle\} \qquad (18)$$

where the function $\mathrm{sign}(.)$ operating on every element of the matrix is

$$\mathrm{sign}(\mathbf{Z}) \triangleq \begin{cases} \frac{Z_{i,j}}{|Z_{i,j}|} & \text{if} \quad Z_{i,j} \neq 0 \\ 0 & \text{if} \quad Z_{i,j} = 0 \end{cases} \qquad (19)$$

In order to clarify how the Algorithm 1 is derived, let us first define the matrix subspaces Ω_X and $\Omega_X^\perp$ as

$$\Omega_X = \{\mathbf{Z} \in \mathbb{C}^{LN \times LN} | Z_{i,j} = 0 \text{ if } X_{i,j} = 0\}$$
$$\Omega_X^\perp = \{\mathbf{Z} \in \mathbb{C}^{LN \times LN} | Z_{i,j} = 0 \text{ if } X_{i,j} \neq 0\}$$

and let $\mathbf{Z}_{\Omega_X}$ and $\mathbf{Z}_{\Omega_X^\perp}$ indicate the projections of matrix $\mathbf{Z}$ onto Ω_X and $\Omega_X^\perp$ respectively.

Theorem 1: For a given $\mathbf{x} = [\mathbf{x}_1' \ldots \mathbf{x}_L']' \in \mathbb{C}^{LN}$, $\mathbf{X} \triangleq \mathbf{x}\mathbf{x}'$ having the eigen-decomposition $\mathbf{X} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}'$, the result $\hat{\mathbf{X}}$ of the optimization ***P-Cal*** is equal to $\mathbf{X}$ **if and only if** all of the following conditions are satisfied:

$$\textbf{C1: if } \mathcal{G}_1 < 0, \text{ then } \lambda < -\frac{1}{\mathcal{G}_1} \qquad (20)$$

$$\textbf{C2: } \mathcal{G}_{-1} > 0 \qquad (21)$$

$$\textbf{C3: } \lambda > \frac{1}{\mathcal{G}_{-1}} \qquad (22)$$

$$\textbf{C4: } \mathcal{G}_0 > 0 \qquad (23)$$

where $\mathcal{G}_p = \mathrm{G}(\mathbf{D}_p)$ and $\mathbf{D}_p$ is defined as

$$\begin{aligned} \mathbf{D}_p \triangleq \arg\min_{\mathbf{Z}} \quad & \mathrm{G}(\mathbf{Z}) && (24)\\ \text{subject to} \quad & \mathbf{m}_i' \mathbf{Z}_{k,\ell} \mathbf{m}_i = 0, && i = 1, \ldots, M \\ & \mathrm{Tr}(\mathbf{Z}) = p, && k, \ell = 1, \ldots, L \\ & \mathbf{Z} = \mathbf{E} \begin{bmatrix} a & \mathbf{b}' \\ \mathbf{b} & \mathbf{C} \end{bmatrix} \mathbf{E}', && a \in \mathbb{R},\ \mathbf{b} \in \mathsf{R}(\mathbf{C}) \\ & \mathbf{C} \succcurlyeq 0, && \mathbf{C} \in \mathbb{C}^{LN-1 \times LN-1} \end{aligned}$$

given that $\mathsf{R}(\mathbf{C})$ represents the range of matrix $\mathbf{C}$ and

$$\mathbf{Z}_{k,\ell} \triangleq \begin{bmatrix} Z_{(k-1)N+1,(\ell-1)N+1} & \cdots & Z_{(k-1)N+1,\ell N} \\ \vdots & \ddots & \vdots \\ Z_{kN,(\ell-1)N+1} & \cdots & Z_{kN,\ell N} \end{bmatrix} \qquad (25)$$

For proof of the theorem, please check the Appendix.

Remark 1: The optimization problem defined in (24) for a given set of $\mathbf{m}_i$, the transform $\mathbf{E}$ and the constant $p \in \mathbb{R}$ is difficult to handle due to the non-linear nature of the constraints specifically introduced by the requirement $\mathbf{b} \in \mathsf{R}(\mathbf{C})$. In order to simplify the optimization, one can omit this criteria and instead solve

Defining the sets $S_{\mathbf{D}}$ and $S_{\hat{\mathbf{D}}}$ as

$$S_{\mathbf{D}} = \left\{ \mathbf{A} | \mathbf{A} = \mathbf{E} \begin{bmatrix} a & \mathbf{b}' \\ \mathbf{b} & \mathbf{C} \end{bmatrix} \mathbf{E}', \quad \begin{matrix} a \in \mathbb{R}, \mathbf{b} \in \mathsf{R}(\mathbf{C}), \mathbf{C} \succcurlyeq 0 \\ \mathbf{C} \in \mathbb{C}^{LN-1 \times LN-1} \end{matrix} \right\} \qquad (26)$$

$$S_{\hat{\mathbf{D}}} = \left\{ \mathbf{A} | \mathbf{A} = \mathbf{E} \begin{bmatrix} a & \mathbf{b}' \\ \mathbf{b} & \mathbf{C} \end{bmatrix} \mathbf{E}', \quad \begin{matrix} a \in \mathbb{R}, \mathbf{b} \in \mathbb{C}^{LN-1}, \mathbf{C} \succcurlyeq 0 \\ \mathbf{C} \in \mathbb{C}^{LN-1 \times LN-1} \end{matrix} \right\} \qquad (27)$$

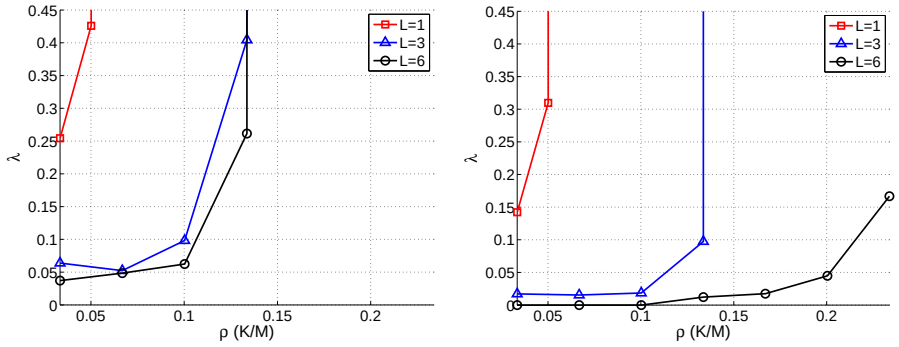
we can observe that $\mathbf{D}_p \in S_{\mathbf{D}}$, $\hat{\mathbf{D}}_p \in S_{\hat{\mathbf{D}}}$ and $S_{\mathbf{D}} \subset S_{\hat{\mathbf{D}}}$. As a result, we can conclude that

1) $\mathrm{G}(\hat{\mathbf{D}}_p) = \mathrm{G}(\mathbf{D}_p)$ if $\hat{\mathbf{D}}_p \in S_{\mathbf{D}}$
2) If $\hat{\mathbf{D}}_p \notin S_{\mathbf{D}}$, then $\mathrm{G}(\hat{\mathbf{D}}_p) \leq \mathrm{G}(\mathbf{D}_p)$ and the bounds on λ computed through Theorem 1 with $\hat{\mathbf{D}}_p$ can only be tighter than or equal to that of the bounds obtained with $\mathbf{D}(p)$.

Following the Theorem 1 and Remark 1, it is straightforward to show that for a given set of sparse input signals and the measurement matrix, Algorithm 1 (***P-Cal-λ***) can be used to determine whether perfect recovery is possible as well as the upper and lower bounds on the parameter λ.

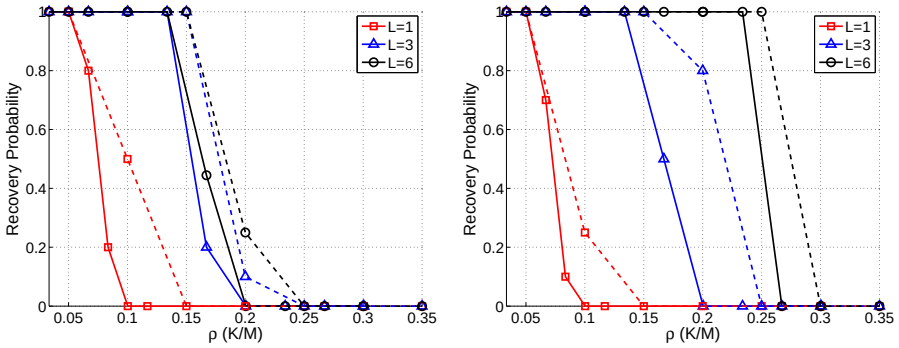
B. Experimental Results for Joint Phase Calibration

In order to test the performance of the proposed approach, phase transition diagrams as in the compressed sensing recovery are plotted for a signal size $N = 100$ with the



(a) Lower bound for λ, $\delta = 0.6$

(b) Lower bound for λ, $\delta = 1.2$



(c) Probability of recovery, $\delta = 0.6$

(d) Probability of recovery, $\delta = 1.2$

Fig. 3: (***P-Cal-λ***) The lower bound on λ, (a)-(b), and the estimated probability of perfect recovery, (c)-(d), for $N = 100$ with respect to $\rho \triangleq K/M$ and L. The **solid lines** indicate the results obtained through ***P-Cal-λ*** whereas **dashed lines** in (c)-(d) indicate the empirical probability of recovery presented on Figure 4 obtained by ***P-Cal***.

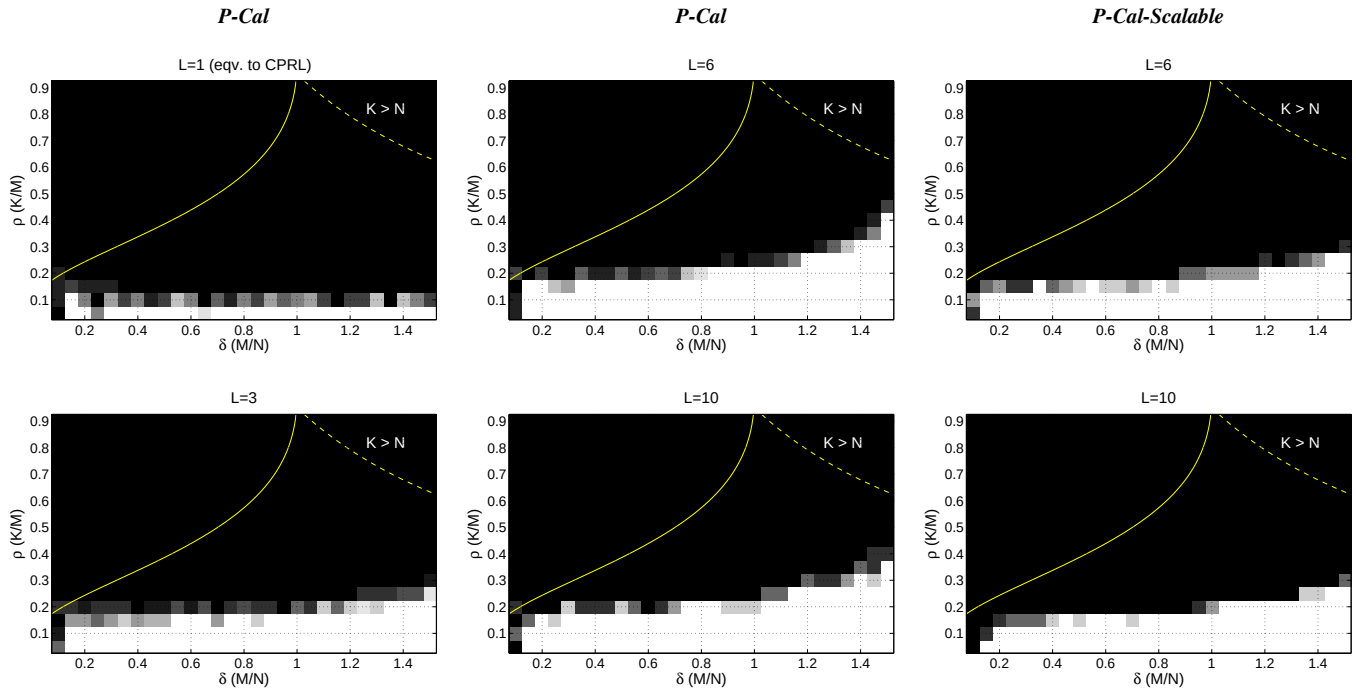


Fig. 4: (***P-Cal*** and ***P-Cal-Scalable***) The probability of perfect recovery (colors black to white indicate probabilities 0 to 1) for $N = 100$ with respect to $\delta \triangleq M/N$ and $\rho \triangleq K/M$. The solid yellow line indicates the Donoho-Tanner phase transition curve for fully calibrated compressed sensing recovery [20]. The dashed yellow line indicates the boundary to the unfeasible region where $K > N$. The rightmost column shows the results with ***P-Cal-Scalable***.

measurement vectors, $\mathbf{m}_i$, and all the non zero entries in the input signals, $\mathbf{x}_\ell$, randomly generated from an i.i.d. normal distribution.

1) Bounds on λ: In order to observe the bounds on λ for perfect recovery, the optimization in (17) is performed for $p = 1, -1, 0$ and the bounds on λ are computed as described in Algorithm 1 (***P-Cal-λ***) for each of the 10 input signals, $\mathbf{x}$, generated randomly as in the earlier experiments. Lowest upper bound and highest lower bound are selected among these 10 experiments as the viable range for λ for a given ρ and δ. For the significant majority of the simulations, λ is found to have no upper bound ($\mathcal{G}_1 > 0$), and therefore the upper bounds are not shown. The maximum lower bound for λ among these 10 experiments as a function of ρ are shown in Figures 3a-3b for $\delta = 0.6$ and $\delta = 1.2$ respectively. It can be observed from Figures 3a and 3b that the benefit of increasing the number of input signals mainly appears when $M > N$ as the recovery is possible for a broader range of λ and ρ. The percentage of recovered signals during these simulations is also displayed along with a comparison with the results predicted with ***P-Cal*** in Figures 3c-3d which verify the consistency among the different simulations. For most of the simulations that resulted in a feasible range of λ for perfect recovery, the optimization result, $\hat{\mathbf{D}}_p$, is observed to be in $S_{\mathbf{D}}$, which affirms that the bounds found on λ for each simulation are tight.

2) Phase Transition Diagrams: The signals (and the phase shift parameters) are recovered for the number of inputs $L = 1, 3, 6, 10$ with the proposed optimization ***P-Cal*** using an ADMM [21] algorithm. The perfect reconstruction criteria is selected as $\frac{1}{L}\sum_\ell \mu(\mathbf{x}_\ell, \hat{\mathbf{x}}_\ell) > 0.95$, where the absolute correlation factor $\mu(\cdot,\cdot)$ is defined as in (10). The threshold for the correlation factor is chosen to be lower than the threshold used for ***A-Cal*** in Section III-A due to ***P-Cal*** being more computationally intensive and hence requiring significantly longer time to reach higher precision. The parameter λ is selected to be within the bounds suggested by simulation results obtained through the algorithm ***P-Cal-λ***.

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 4 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 [15]. It can be observed that the proposed joint recovery methods provide significantly better performance than individual optimization, 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.

C. Discussion

The are several remarks that can be made on the performance of joint phase calibration with ***P-Cal***

1) The proposed method 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 for all θ_i and $\mathbf{x}_l$ unlike when $\mathbf{x}_l$ are recovered by individual optimization as in (5).

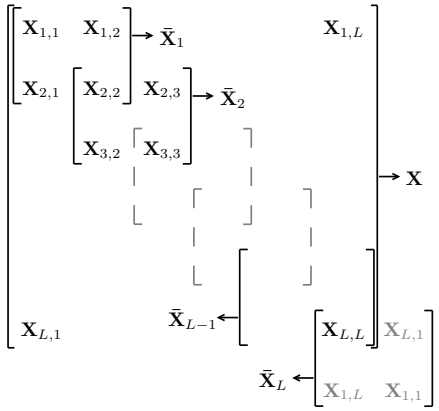


Fig. 5: The sub-matrices $\bar{\mathbf{X}}_1, \ldots, \bar{\mathbf{X}}_L$ within the matrix $\mathbf{X}$.

2) Even though the optimization method utilized in (15) is similar to (5), 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.
3) The empirical performance suggests that only few input signals are sufficient for ***P-Cal*** when $M < N$. Even though more input signals provide improvement in the performance for $M > N$, this case can also be handled by obtaining a closed form solution (or ***LS-Cal***) with a large enough L as discussed in Section III. However when there is also a variation in the amplitude of the gains, a larger number of input signals may be required which will be discussed in Section VI.
4) 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 as discussed in [11]. However the optimization is still challenging for large L. A modified version of ***P-Cal*** is presented in Section V in order to provide a more scalable performance with respect to L.

V. 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 ***P-Cal***. To address this problem, an alternative approach is to construct sub-matrices of $\mathbf{X}$:

$$\bar{\mathbf{X}}_\ell \triangleq \begin{bmatrix} \mathbf{X}_{\ell,\ell} & \mathbf{X}_{\ell,(\ell)_L+1} \\ \mathbf{X}_{(\ell)_L+1,\ell} & \mathbf{X}_{(\ell)_L+1,(\ell)_L+1} \end{bmatrix} \quad \ell = 1, \ldots, L \tag{28}$$

where $(\ell)_L \triangleq \ell \pmod L$. A depiction of the sub-matrices with respect to $\mathbf{X}$ can be seen in Figure 5. Similarly to $\mathbf{X}$, $\bar{\mathbf{X}}_\ell$ are also rank-one, hermitian, positive semi-definite and sparse matrices. Therefore, an alternative to the optimization ***P-Cal***

TABLE I: 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
P-Cal	ML^2	$NL \times NL$ matrix	L signals with single global phase shift
P-Cal-Scalable	$3ML$	L matrices of $2N \times 2N$	L signals with single global phase shift

is solving

***P-Cal-Scalable*:**

$$\hat{\bar{\mathbf{X}}}_1, \ldots, \hat{\bar{\mathbf{X}}}_L = \underset{\bar{\mathbf{Z}}_1 \ldots \bar{\mathbf{Z}}_L}{\arg\min} \sum_{k=1}^{L} f_\lambda(\bar{\mathbf{Z}}_k) \tag{29}$$

$$\text{subject to} \quad \bar{\mathbf{Z}}_1 \succcurlyeq 0, \ldots, \bar{\mathbf{Z}}_L \succcurlyeq 0$$

$$i = 1, \ldots, M \quad g_{i,\ell,\ell} = \mathbf{m}_i' \mathbf{Z}_{\ell,\ell} \mathbf{m}_i$$

$$\ell = 1, \ldots, L \quad g_{i,(\ell)_L+1,\ell} = \mathbf{m}_i' \mathbf{Z}_{(\ell)_L+1,\ell} \mathbf{m}_i$$

$$g_{i,\ell,(\ell)_L+1} = \mathbf{m}_i' \mathbf{Z}_{\ell,(\ell)_L+1} \mathbf{m}_i$$

$$\bar{\mathbf{Z}}_\ell = \begin{bmatrix} \mathbf{Z}_{\ell,\ell} & \mathbf{Z}_{\ell,(\ell)_L+1} \\ \mathbf{Z}_{(\ell)_L+1,\ell} & \mathbf{Z}_{(\ell)_L+1,(\ell)_L+1} \end{bmatrix}$$

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

$$\hat{\bar{\mathbf{X}}}_\ell = \begin{bmatrix} \hat{\mathbf{x}}_\ell \\ \hat{\mathbf{x}}_{(\ell)_L+1} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_\ell' & \hat{\mathbf{x}}_{(\ell)_L+1}' \end{bmatrix} \quad \ell = 1 \ldots L \tag{30}$$

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

The optimization ***P-Cal-Scalable*** deals with L matrices of size $2N \times 2N$ instead of a single $LN \times LN$ matrix. The number of constraints is also reduced from ML^2 to $3ML$ for $L > 3$ with respect to ***P-Cal*** while retaining the joint recovery characteristic. A comparison can be seen in Table I. The matrices $\bar{\mathbf{X}}_1, \ldots, \bar{\mathbf{X}}_L$ have comparable size to individual recovery as in the CPRL or Phaselift algorithms.

A. Experimental Results for Scalable Phase Calibration

The performance of ***P-Cal-Scalable*** can be seen in the rightmost column in Figure 4. Although there is a slight performance degradation with respect to ***P-Cal***, 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 sub-matrix 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 require more iterations due to larger number of auxiliary variables in ADMM implementation ***P-Cal-Scalable***.

B. Discussion

Despite the fact that the ***P-Cal*** is based on Quadratic Basis Pursuit optimization like CPRL, the measurement constraints for the joint calibration problem has a separable and significantly overcomplete structure unlike CPRL and the phase retrieval problem. This separability enables the formulation of the problem through the recovery of the sub-matrices as in ***P-Cal-Scalable*** while performing recovery by enforcing only a subset of the measurement constraints.

In this framework, the scalable approach in ***P-Cal-Scalable*** is beneficial when a large number of L is required for recovery. Even though a small L is shown to be sufficient for phase calibration in Section IV, larger number of input signals may be needed when there is variation among the gain amplitudes as well. This case will be discussed in more detail in Section VI in which the approach used in ***P-Cal-Scalable*** will be further developed.

VI. Complete Calibration

In order to perform the joint recovery of the gains and the sparse inputs when the amplitudes as well as the phases of the sensor gains vary significantly, we can combine the amplitude and phase calibration approaches introduced in Sections III-V. The cross measurements are equal to

$$g_{i,k,\ell} \triangleq y_{i,k} y_{i,\ell}' \qquad i = 1 \ldots M,\ k, \ell = 1 \ldots L \tag{31}$$

$$= d_i e^{j\theta_i} \mathbf{m}_i' \mathbf{x}_k \mathbf{x}_\ell' \mathbf{m}_i e^{-j\theta_i} d_i \tag{32}$$

$$= d_i^2 \mathbf{m}_i' \mathbf{X}_{k,\ell} \mathbf{m}_i \quad d_i \in \mathbb{R} \tag{33}$$

which in turn can be used in the combined optimization

***C-Cal*:**

$$\underset{\hat{\tau}_1, \ldots, \hat{\tau}_M}{\hat{\mathbf{X}}} = \underset{\mathbf{Z}, t_1, \ldots, t_M}{\arg\min} f_\lambda(\mathbf{Z}) \tag{34}$$

$$\text{subject to } t_i g_{i,k,\ell} = \mathbf{m}_i' \mathbf{Z}_{k,\ell} \mathbf{m}_i,\ i = 1 \ldots M$$

$$\mathbf{Z} \succcurlyeq 0, \qquad k, \ell = 1 \ldots L$$

$$\sum_{n=1}^{M} t_n = c$$

The unknown gains can be set as $\hat{d}_i = \frac{1}{\sqrt{\hat{\tau}_i}}$ and the unknown sparse signals can be recovered from $\hat{\mathbf{X}}$ the same as in the phase calibration optimization ***P-Cal***.

Unlike the phase calibration, a large number of input signals may be required in complete calibration when the amplitude variation of the sensor gains are high. Therefore the scalable approach in ***P-Cal-Scalable*** can provide a significant improvement in the complexity and memory requirements.

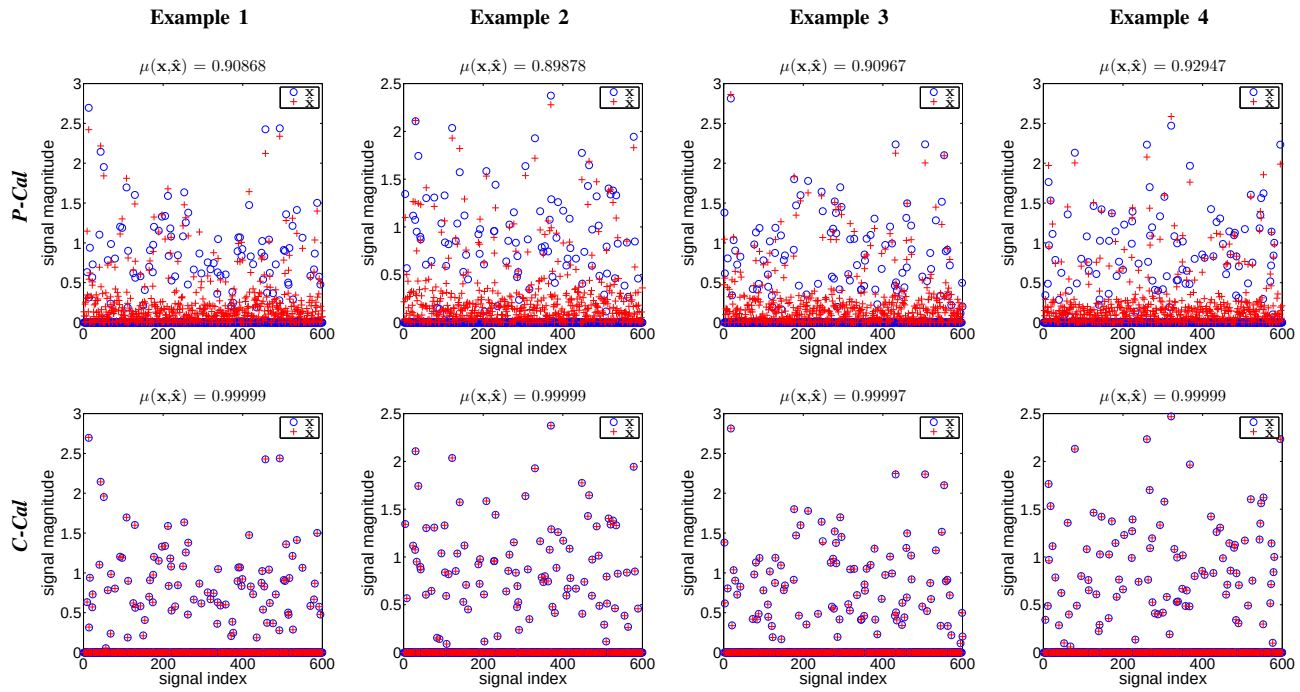


Fig. 6: (***C-Cal***) Different example signals recovered with ***P-Cal*** (top row) and ***C-Cal*** (bottom row) for $N = 100$, $L = 6$, $\delta = 0.8$, $\rho = 0.2$, $\sigma = 0.2$. The magnitude of the original signal, $\mathbf{x}$, and the estimated signal, $\hat{\mathbf{x}}$, after scale normalization are plotted for comparison. The correlation factor between the original and the estimated signals are also reported above each figure.

The scalable version of the ***C-Cal*** can be formulated as

***C-Cal-Scalable*:**

$$\begin{aligned}
\underset{\hat{\tau}_1,\ldots,\hat{\tau}_M}{\hat{\bar{\mathbf{X}}}_1,\ldots,\hat{\bar{\mathbf{X}}}_L} &= \underset{\substack{\bar{\mathbf{Z}}_1\ldots\bar{\mathbf{Z}}_L \\ t_1,\ldots,t_M}}{\arg\min} \quad \sum_{k=1}^{L} f_\lambda(\bar{\mathbf{Z}}_k) \qquad (35)\\
&\text{subject to} \quad \bar{\mathbf{Z}}_1 \succcurlyeq 0,\ldots,\bar{\mathbf{Z}}_L \succcurlyeq 0\\
&i = 1\ldots M \quad t_i g_{i,\ell,\ell} = \mathbf{m}_i' \mathbf{Z}_{\ell,\ell} \mathbf{m}_i\\
&\ell = 1\ldots L \quad t_i g_{i,(\ell)_L+1,\ell} = \mathbf{m}_i' \mathbf{Z}_{(\ell)_L+1,\ell} \mathbf{m}_i\\
&\qquad\quad t_i g_{i,\ell,(\ell)_L+1} = \mathbf{m}_i' \mathbf{Z}_{\ell,(\ell)_L+1} \mathbf{m}_i\\
&\qquad\quad \bar{\mathbf{Z}}_\ell = \begin{bmatrix} \mathbf{Z}_{\ell,\ell} & \mathbf{Z}_{\ell,(\ell)_L+1} \\ \mathbf{Z}_{(\ell)_L+1,\ell} & \mathbf{Z}_{(\ell)_L+1,(\ell)_L+1} \end{bmatrix}\\
&\qquad\quad \sum_{n=1}^{M} t_n = c
\end{aligned}$$

which is simply the combination of the optimization approaches in ***A-Cal*** and ***P-Cal-Scalable***.

A. Experimental Results for Complete Calibration

The optimization methods ***C-Cal*** and ***C-Cal-Scalable*** for complete calibration are combinations of the optimization methods provided in Sections III-V, therefore a detailed analysis of these methods will not be provided. Instead a few sample recovery results are shown in Figure 6 to demonstrate the performance difference between using phase calibration and complete calibration in a system where the amplitudes of the sensor gains are randomly generated as described in Section III-A with $\sigma = 0.2$. These examples demonstrate that both ***A-Cal*** and ***P-Cal*** fail when both the amplitudes and the phases of the unknown gains vary significantly and that the combined approach introduced in the complete calibration is required for recovery.

B. Discussion

The complete calibration approach ***C-Cal*** is the combination of amplitude and phase calibration approaches introduced in earlier sections and therefore its performance can be expected to have similar characteristics to ***A-Cal*** and ***P-Cal***. Since it is an extension of the quadratic formulation used in ***P-Cal***, the complete calibration also has similar limitations in performance for low sparsity (large ρ) as well as with respect to memory and complexity. As observed with ***A-Cal***, it is possible to calibrate very large variations (large σ) in amplitude of the gains, provided that the number of input signals, L, is sufficiently high. When many input signals are needed, the scalable approach ***C-Cal-Scalable*** can be utilized to take advantage of the memory requirements and complexity that are linear in L.

VII. Conclusions

Several convex optimization methods to handle blind calibration of complex valued gains in sparse inverse problems has been proposed in this paper[1]. Investigating the performance of all the presented methods, it can be seen that all

[1]The codes for the MATLAB® implementations of the proposed methods has been provided in http://hal.inria.fr/docs/00/85/32/25/ANNEX/Calcodesv1.0.rar

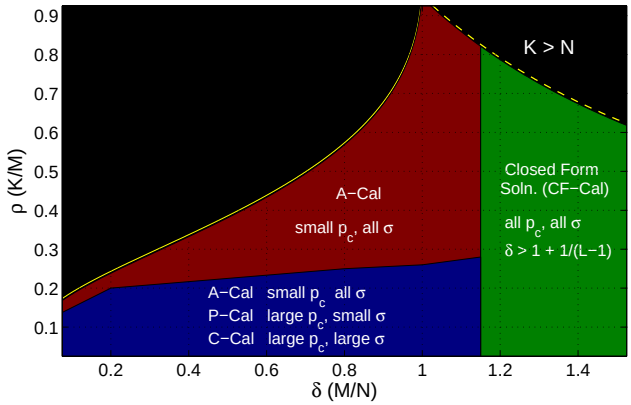


Fig. 7: A depiction of the algorithms to be used for different calibration scenarios. Note that for each case L must also be sufficiently high depending on σ and p_c. The solid yellow line indicates the Donoho-Tanner phase transition curve for fully calibrated compressed sensing recovery [20].

the proposed optimization methods provide complementary advantages and each can be used in a different scenario depending on the distribution of the unknown sensor gains. Even though complete calibration can be used for all cases, the use of amplitude calibration or phase calibration may be preferable for a simplified optimization when possible. A summary of which algorithm to use for different conditions on the sensor gains are shown in Figure 7 along with the successful recovery regions of the algorithms in the $\delta - \rho$ plane. When a large number, L, of input signals is required for recovery in phase or complete calibration, the scalable versions of the algorithms can be used for less memory requirements and less complexity during reconstruction. Note that recovery in the red region of Figure 7 is only possible for small variation in phases of the sensor gains due to limitations in the performance of calibration methods with quadratic formulation (***P-Cal***, ***P-Cal-Scalable***, ***C-Cal***, ***C-Cal-Scalable***).

It is observed for ***P-Cal*** that the parameter λ has no upper bound for majority of the randomly generated signals which implies that recovery performance is maximized for majority of signals if λ is selected to be sufficiently large. A more interesting implication of this result is that the $\mathrm{Tr}(.)$ in ***P-Cal*** optimization may not be needed for perfect recovery for majority of the signals. This is also consistent with the deductions in [23] showing that the combination of $\mathrm{Tr}(.)$ and $\|.\|_1$ in the objective function for simultaneously enforcing low-rank and sparsity can only be as effective as using one of them. We aim to further investigate this aspect of the optimization on low-rank and sparse matrices.

It is possible to overcome the limitations in the performance of the proposed quadratic methods through efficient algorithms and non-convex methods which are planned to be investigated as future work. The performance under noise as well as calibration of real world data from different application fields is also another direction for planned research.

APPENDIX
PROOF OF THEOREM 1

In order to prove Theorem 1 we shall first establish a few observations. Let us define the cone S_X such that

$$S_X = \{\mathbf{A} | \mathbf{X} + c\mathbf{A} \succcurlyeq 0,\ \exists c > 0\} \tag{A.1}$$

Lemma 1: For a given $\mathbf{x} = [\mathbf{x}_1' \dots \mathbf{x}_L']' \in \mathbb{C}^{LN}$, $\mathbf{X} = \mathbf{x}\mathbf{x}'$ having the eigen-decomposition $\mathbf{X} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}'$, the matrix $\mathbf{\Delta} \triangleq \mathbf{E}\begin{bmatrix} a & \mathbf{b}' \\ \mathbf{b} & \mathbf{C} \end{bmatrix}\mathbf{E}'$ where $\mathbf{b} \in \mathbb{C}^{LN-1}$, $\mathbf{C} \in \mathbb{C}^{LN-1 \times LN-1}$ is in S_X **if and only if** $\mathbf{C} \succcurlyeq 0$, $a \in \mathbb{R}$ and $\mathbf{b} \in \mathsf{R}(\mathbf{C})$ where $\mathsf{R}(.)$ represents the range of the matrix.

Proof of Lemma 1: Let us assume that $\mathbf{\Delta} \in S_X$, then by definition

$$\exists c_0 \in \mathbb{R}^+, \text{ such that } \forall u \in \mathbb{C}, \forall \mathbf{v} \in \mathbb{C}^{LN-1}, \forall c \in (0, c_0],$$

$$[u'\ \mathbf{v}'] \left(\mathbf{X} + c\mathbf{E}\begin{bmatrix} a & \mathbf{b}' \\ \mathbf{b} & \mathbf{C} \end{bmatrix}\mathbf{E}' \right) \begin{bmatrix} u \\ \mathbf{v} \end{bmatrix} \geq 0 \tag{A.2}$$

$$\Rightarrow [u'\ \mathbf{v}'] \left(\mathbf{\Delta} + c\begin{bmatrix} a & \mathbf{b}' \\ \mathbf{b} & \mathbf{C} \end{bmatrix} \right) \begin{bmatrix} u \\ \mathbf{v} \end{bmatrix} \geq 0 \tag{A.3}$$

$$\Rightarrow [u'\ \mathbf{v}'] \begin{bmatrix} \|\mathbf{x}\|_2^2 + ca & c\mathbf{b}' \\ c\mathbf{b} & c\mathbf{C} \end{bmatrix} \begin{bmatrix} u \\ \mathbf{v} \end{bmatrix} \geq 0 \tag{A.4}$$

$$\Rightarrow |u|^2(\|\mathbf{x}\|_2^2 + ca) + cu'\mathbf{b}'\mathbf{v} + cu\mathbf{v}'\mathbf{b} + c\mathbf{v}'\mathbf{C}\mathbf{v} \geq 0 \tag{A.5}$$

The first necessary condition for (A.5) is that

$$\|\mathbf{x}\|_2^2 + ca \geq 0 \quad \forall c \in (0, c_0] \Rightarrow a \geq \frac{-\|\mathbf{x}\|_2^2}{c_0}, \quad a \in \mathbb{R} \tag{A.6}$$

Following (A.5), we have

$$(\|\mathbf{x}\|_2^2 + ca)\left[\left\| u + \frac{c\mathbf{b}'\mathbf{v}}{\|\mathbf{x}\|_2^2 + ca} \right\|_2^2 - \left\| \frac{c\mathbf{b}'\mathbf{v}}{\|\mathbf{x}\|_2^2 + ca} \right\|_2^2 \right] + c\mathbf{v}'\mathbf{C}\mathbf{v} \geq 0 \tag{A.7}$$

$$\Rightarrow \mathbf{v}'\mathbf{C}\mathbf{v} - \frac{c\mathbf{v}'\mathbf{b}\mathbf{b}'\mathbf{v}}{\|\mathbf{x}\|_2^2 + ca} \geq 0, \quad \forall \mathbf{v},\ \forall c \in (0, c_0] \tag{A.8}$$

$$\Rightarrow \mathbf{C} - \frac{c\mathbf{b}\mathbf{b}'}{\|\mathbf{x}\|_2^2 + ca} \succcurlyeq 0, \quad \forall c \in (0, c_0] \tag{A.9}$$

The second and third necessary conditions implied by (A.9) are

$$\mathbf{C} \succcurlyeq 0 \tag{A.10}$$

$$\mathbf{b}'\mathbf{v} = 0 \quad \forall \mathbf{v} \text{ satisfying } \mathbf{C}\mathbf{v} = 0 \quad \Rightarrow \quad \mathbf{b} \in \mathsf{R}(\mathbf{C}) \tag{A.11}$$

More strict conditions can be derived assuming $\mathbf{C}$ has the eigen-decomposition $\mathbf{C} = \mathbf{F}\mathbf{\Lambda}_\mathbf{C}\mathbf{\Lambda}_\mathbf{C}'\mathbf{F}'$, $\mathbf{\Lambda}_\mathbf{C} = \mathrm{Diag}(\sqrt{\lambda_{\mathbf{C},1}}, \dots, \sqrt{\lambda_{\mathbf{C},LN-1}})$ and without loss of generality representing $\mathbf{b}$ as

$$\mathbf{b} = \mathbf{F}\mathbf{\Lambda}_\mathbf{C}\mathbf{s}, \qquad \mathbf{s} \in \mathbb{C}^{LN-1} \tag{A.12}$$

$$= \mathbf{F}\mathbf{\Lambda}_\mathbf{C}\mathbf{s}_{\Omega_\mathbf{C}}, \quad \Omega_\mathbf{C} \triangleq \{\mathbf{a} | a_i = 0 \quad \text{if} \quad \lambda_{\mathbf{C},i} = 0\} \tag{A.13}$$

$$\Rightarrow \mathbf{C} - \frac{c\mathbf{b}\mathbf{b}'}{\|\mathbf{x}\|_2^2 + ca} = \mathbf{F}\mathbf{\Lambda}_\mathbf{C}\mathbf{\Lambda}_\mathbf{C}'\mathbf{F}' - \frac{c\mathbf{F}\mathbf{\Lambda}_\mathbf{C}\mathbf{s}_{\Omega_\mathbf{C}}\mathbf{s}_{\Omega_\mathbf{C}}'\mathbf{\Lambda}_\mathbf{C}'\mathbf{F}'}{\|\mathbf{x}\|_2^2 + ca} \succcurlyeq 0 \tag{A.14}$$

$$\Rightarrow \hat{\mathbf{v}}'\hat{\mathbf{v}} - \frac{c}{\|\mathbf{x}\|_2^2 + ca}|\hat{\mathbf{v}}'\mathbf{s}_{\Omega_\mathbf{C}}|^2 \geq 0, \quad \forall \mathbf{v} \in \mathbb{C}^{LN-1} \quad \text{(A.15)}$$

$$\hat{\mathbf{v}} \triangleq \mathbf{\Lambda}'_\mathbf{C}\mathbf{F}'\mathbf{v}$$

$$\Rightarrow 1 - \frac{c}{\|\mathbf{x}\|_2^2 + ca}\left|\frac{\hat{\mathbf{v}}'\mathbf{s}_{\Omega_\mathbf{C}}}{|\hat{\mathbf{v}}|}\right|^2 \geq 0, \quad \forall \mathbf{v} \in \mathbb{C}^{LN-1} \quad \text{(A.16)}$$

$$\Rightarrow 1 - \frac{c}{\|\mathbf{x}\|_2^2 + ca}\|\mathbf{s}_{\Omega_\mathbf{C}}\|_2^2 \geq 0 \quad \text{(A.17)}$$

$$\Rightarrow \|\mathbf{s}_{\Omega_\mathbf{C}}\|_2^2 \leq a + \frac{\|\mathbf{x}\|_2^2}{c} \quad \forall c \in (0, c_0] \quad \text{(A.18)}$$

Given c_0, the three necessary conditions are $a \in \mathbb{R}$, $\mathbf{C} \succcurlyeq 0$ and $\mathbf{b} \in \mathsf{R}(\mathbf{C})$ as shown in (A.6), (A.10) and (A.11). It can also be shown that given $\mathbf{\Delta}$, a constant, c_0, can be chosen to ensure $\mathbf{\Delta} \in S_X$ considering the conditions in (A.6) and (A.18) provided that $a \in \mathbb{R}$, $\mathbf{C} \succcurlyeq 0$ and $\mathbf{b} \in \mathsf{R}(\mathbf{C})$. ■

Lemma 2: The matrix $\mathbf{X}$ is not the global (or local) minimum of the optimization ***P-Cal*** **if and only if** $\exists \mathbf{\Delta} \in \mathbb{C}^{LN \times LN}$ satisfying all three of the following conditions:

$$\textbf{C1: } \mathbf{m}_i'\mathbf{\Delta}_{k,\ell}\mathbf{m}_i = 0,\ i = 1, \ldots, M,\ k, \ell = 1, \ldots, L \quad \text{(A.19)}$$

$$\textbf{C2: } \mathbf{\Delta} \in S_X \quad \text{(A.20)}$$

$$\textbf{C3: } \mathrm{Tr}(\mathbf{\Delta}) + \lambda\, \mathrm{G}(\mathbf{\Delta}) \leq 0 \quad \text{(A.21)}$$

where $\mathbf{\Delta}_{k,\ell}$ is defined similar to (25).

Corollary 1: The matrix $\mathbf{X}$ is the global minimum of the optimization ***P-Cal*** **if and only if**

$$\mathrm{Tr}(\mathbf{\Delta}) + \lambda\, \mathrm{G}(\mathbf{\Delta}) > 0,\ \forall \mathbf{\Delta} \text{ satisfying } \textbf{C1} \text{ and } \textbf{C2} \quad \text{(A.22)}$$

Proof of Lemma 2: If $\mathbf{X}$ is not the global minimum of the optimization ***P-Cal***, then by definition $\exists \mathbf{W} \succcurlyeq 0$, $\mathbf{W} \neq \mathbf{X}$ such that

$$f_\lambda(\mathbf{W}) \leq f_\lambda(\mathbf{X}) \quad \text{(A.23)}$$

$$g_{i,k,\ell} = \mathbf{m}_i'\mathbf{W}_{k,\ell}\mathbf{m}_i,\ i = 1, \ldots, M,\ k, \ell = 1, \ldots, L \quad \text{(A.24)}$$

Using (A.23) and convexity of the function f_λ

$$f_\lambda(\mathbf{W}) \leq f_\lambda(\mathbf{X} + c\mathbf{\Delta}) \leq f_\lambda(\mathbf{X}), \quad 0 < c \leq 1 \quad \text{(A.25)}$$

$$\mathbf{\Delta} \triangleq \mathbf{W} - \mathbf{X}$$

Considering that $\mathbf{X}$ satisfies the measurements ($g_{i,k,\ell} = \mathbf{m}_i'\mathbf{X}_{k,\ell}\mathbf{m}_i,\ i = 1, \ldots, M,\ k, \ell = 1, \ldots, L$), (A.24) easily leads us to **C1** such that

$$\mathbf{m}_i'(\mathbf{W} - \mathbf{X})_{k,\ell}\mathbf{m}_i = 0,\ i = 1, \ldots, M,\ k, \ell = 1, \ldots, L \quad \text{(A.26)}$$

$$\Rightarrow \mathbf{m}_i'\mathbf{\Delta}_{k,\ell}\mathbf{m}_i = 0,\ 0 < c \leq 1 \quad \text{(A.27)}$$

Note that (A.25), (A.27) and the fact that $\mathbf{X} + c\mathbf{\Delta} \succcurlyeq 0$, $0 < c \leq 1$ due to convexity of the space of positive semi-definite matrices easily implies $\mathbf{X}$ not being a local minimum when a global minimum $\mathbf{W}$ exists as expected by the convexity of the optimization ***P-Cal***. $\mathbf{\Delta} \in S_X$ (**C2**) is also implied by definition given that $\mathbf{X} + c\mathbf{\Delta} \succcurlyeq 0$, $0 < c \leq 1$.

Continuing from (A.25)

$$f_\lambda(\mathbf{X} + c\mathbf{\Delta}) \leq f_\lambda(\mathbf{X}) \quad \text{(A.28)}$$

$$\Rightarrow \lim_{c \to 0^+} \frac{f_\lambda(\mathbf{X} + c\mathbf{\Delta}) - f_\lambda(\mathbf{X})}{c} \leq 0, \qquad 0 < c \leq 1 \quad \text{(A.29)}$$

$$\Rightarrow \lim_{c \to 0^+} \frac{1}{c}\Big[\, \mathrm{Tr}(\mathbf{X} + c\mathbf{\Delta}) + \lambda\|\mathbf{X} + c\mathbf{\Delta}\|_1 - \mathrm{Tr}(\mathbf{X}) - \lambda\|\mathbf{X}\|_1\Big] \leq 0 \quad \text{(A.30)}$$

$$\Rightarrow \lim_{c \to 0^+} \frac{1}{c}\Big[c\,\mathrm{Tr}(\mathbf{\Delta}) + \lambda\langle \mathrm{sign}(\mathbf{X} + c\mathbf{\Delta}), \mathbf{X} + c\mathbf{\Delta}\rangle - \lambda\|\mathbf{X}\|_1\Big] \leq 0 \quad \text{(A.31)}$$

Note that

$$\langle \mathrm{sign}(\mathbf{X} + c\mathbf{\Delta}), \mathbf{X} + c\mathbf{\Delta}\rangle = \sum_{i,j} \mathrm{sign}(X_{i,j} + c\Delta_{i,j})(X_{i,j} + c\Delta_{i,j}) \quad \text{(A.32)}$$

and for small enough c, each term of (A.32) can be reduced to

$$\mathrm{sign}(X_{i,j} + c\Delta_{i,j})[X_{i,j} + c\Delta_{i,j}] = \begin{cases} c\left|\Delta_{i,j}\,\mathrm{Imag}\left\{\mathrm{sign}(\frac{\Delta_{i,j}}{X_{i,j}})\right\}\phi(X_{i,j}, c\Delta_{i,j})\right| & \\ \quad +|X_{i,j}| + c|\Delta_{i,j}|\,\mathrm{Real}\left\{\mathrm{sign}(\frac{\Delta_{i,j}}{X_{i,j}})\right\} & X_{i,j} \neq 0 \\ c|\Delta_{i,j}| & X_{i,j} = 0 \end{cases} \quad \text{(A.33)}$$

where

$$\phi(X_{i,j}, c\Delta_{i,j}) \triangleq \frac{\mathrm{Imag}\left\{\sqrt{\mathrm{sign}\left(\frac{X_{i,j} + c\Delta_{i,j}}{X_{i,j}}\right)}\right\}}{\mathrm{Real}\left\{\sqrt{\mathrm{sign}\left(\frac{X_{i,j} + c\Delta_{i,j}}{X_{i,j}}\right)}\right\}} \quad \text{(A.34)}$$

Considering

$$\lim_{c \to 0^+} \phi(X_{i,j}, c\Delta_{i,j}) = 0 \quad \text{(A.35)}$$

and

$$|\Delta_{i,j}|\,\mathrm{Real}\left\{\mathrm{sign}\left(\frac{\Delta_{i,j}}{X_{i,j}}\right)\right\} = \mathrm{Real}\left\{\Delta_{i,j}\,\mathrm{sign}(X'_{i,j})\right\} \quad \text{(A.36)}$$

the term in (A.31) then reduces to

$$\lim_{c \to 0^+} \frac{1}{c}\Big[c\,\mathrm{Tr}(\mathbf{\Delta}) + \lambda\langle \mathrm{sign}(\mathbf{X} + c\mathbf{\Delta}), \mathbf{X} + c\mathbf{\Delta}\rangle - \lambda\|\mathbf{X}\|_1\Big] \leq 0 \quad \text{(A.37)}$$

$$\Rightarrow \lim_{c \to 0^+} \frac{1}{c}\Bigg[c\,\mathrm{Tr}(\mathbf{\Delta}) + \lambda\Big[\|\mathbf{X}\|_1 + c\|\mathbf{\Delta}_{\Omega_\mathbf{X}^\perp}\|_1 + c\,\mathrm{Real}\left\{\langle \mathrm{sign}(\mathbf{X}), \mathbf{\Delta}_{\Omega_X}\rangle\right\}\Big] - \lambda\|\mathbf{X}\|_1\Bigg] \leq 0 \quad \text{(A.38)}$$

$$\Rightarrow \lim_{c \to 0^+} \frac{1}{c}\Big[c\,\mathrm{Tr}(\mathbf{\Delta}) + c\lambda\|\mathbf{\Delta}_{\Omega_\mathbf{X}^\perp}\|_1 + c\lambda\,\mathrm{Real}\left\{\langle \mathrm{sign}(\mathbf{X}), \mathbf{\Delta}_{\Omega_X}\rangle\right\}\Big] \leq 0 \quad \text{(A.39)}$$

$$\Rightarrow \mathrm{Tr}(\mathbf{\Delta}) + \lambda\|\mathbf{\Delta}_{\Omega_\mathbf{X}^\perp}\|_1 + \lambda\,\mathrm{Real}\left\{\langle \mathrm{sign}(\mathbf{X}), \mathbf{\Delta}_{\Omega_X}\rangle\right\} = \mathrm{Tr}(\mathbf{\Delta}) + \lambda\,\mathrm{G}(\mathbf{\Delta}) \leq 0 \quad \text{(A.40)}$$

which is exactly **C3**. Hence existence of a global minimum $\mathbf{W} \neq \mathbf{X}$ implies **C1**, **C2** and **C3**. Similarly, it can also easily be shown that conditions **C1**, **C2** and **C3** are sufficient for $\mathbf{X}$ not being the minimum of the convex optimization ***P-Cal***. ■

Proof of Theorem 1: Following the Corollary 1, in order to have $\mathbf{X}$ as a unique solution to optimization ***P-Cal***, we must have

$$\begin{aligned} &\mathrm{Tr}(\mathbf{\Delta}) + \lambda\,\mathrm{G}(\mathbf{\Delta}) > 0 \qquad \text{(A.41)}\\ &\quad \forall \mathbf{\Delta} \text{ satisfying } \mathbf{m}_i'\mathbf{\Delta}_{k,\ell}\mathbf{m}_i = 0, \quad i = 1,\ldots,M\\ &\qquad\qquad \mathbf{\Delta} \in S_X, \qquad k,\ell = 1,\ldots,L \end{aligned}$$

or equivalently

$$\begin{aligned} &\mathrm{Tr}(\mathbf{\Delta}) + \lambda\,\mathrm{G}(\mathbf{\Delta}) > 0 \qquad \text{(A.42)}\\ &\quad \forall \mathbf{\Delta} \text{ satisfying } \mathbf{m}_i'\mathbf{\Delta}_{k,\ell}\mathbf{m}_i = 0,\ i = 1,\ldots,M,\ k,\ell = 1,\ldots,L\\ &\qquad\qquad \mathbf{\Delta} = \mathbf{E}\begin{bmatrix} a & \mathbf{b}'\\ \mathbf{b} & \mathbf{C}\end{bmatrix}\mathbf{E}',\ a \in \mathbb{R},\ \mathbf{b} \in \mathsf{R}(\mathbf{C})\\ &\qquad\qquad \mathbf{C} \succcurlyeq 0, \qquad\qquad \mathbf{C} \in \mathbb{C}^{LN-1\times LN-1} \end{aligned}$$

as suggested by Lemma 1. Since given a $\mathbf{\Delta}$ satisfying (A.42) $c\mathbf{\Delta}$, $c > 0$ also satisfies (A.42), $\mathrm{Tr}(\mathbf{\Delta})$ can be fixed without loss of generality. Therefore (A.42) can be considered in three cases:

- **Case 1:** $\mathrm{Tr}(\mathbf{\Delta}) = 1$
  (A.42) can be satisfied for all $\mathrm{Tr}(\mathbf{\Delta}) > 0$ provided that

  $$\begin{aligned} &\mathrm{Tr}(\mathbf{\Delta}) + \lambda\,\mathrm{G}(\mathbf{\Delta}) > 0 \qquad \text{(A.43)}\\ &\quad \forall \mathbf{\Delta} \text{ s.t. } \mathbf{m}_i'\mathbf{\Delta}_{k,\ell}\mathbf{m}_i = 0,\ i = 1,\ldots,M,\ k,\ell = 1,\ldots,L\\ &\qquad\qquad \mathrm{Tr}(\mathbf{\Delta}) = 1\\ &\qquad\qquad \mathbf{\Delta} = \mathbf{E}\begin{bmatrix} a & \mathbf{b}'\\ \mathbf{b} & \mathbf{C}\end{bmatrix}\mathbf{E}',\ a \in \mathbb{R},\ \mathbf{b} \in \mathsf{R}(\mathbf{C})\\ &\qquad\qquad \mathbf{C} \succcurlyeq 0, \qquad\qquad \mathbf{C} \in \mathbb{C}^{LN-1\times LN-1} \end{aligned}$$

  which is satisfied if $\lambda\,\mathrm{G}(\mathbf{D}_1) > -1$. Therefore if $\mathrm{G}(\mathbf{D}_1) < 0$, we must have $\lambda < \frac{-1}{\mathrm{G}(\mathbf{D}_1)}$, and if $\mathrm{G}(\mathbf{D}_1) \geq 0$ no limitation on λ is needed (**C1**).
- **Case 2:** $\mathrm{Tr}(\mathbf{\Delta}) = -1$
  Similar to Case 1, (A.42) can be satisfied for all $\mathrm{Tr}(\mathbf{\Delta}) < 0$ provided that $\lambda\,\mathrm{G}(\mathbf{D}_{-1}) > 1$. Consequently, $\lambda\,\mathrm{G}(\mathbf{D}_{-1}) > 1$ only if $\mathrm{G}(\mathbf{D}_{-1}) > 0$ (**C2**) and $\lambda > \frac{1}{\mathrm{G}(\mathbf{D}_{-1})}$ (**C3**).
- **Case 3:** $\mathrm{Tr}(\mathbf{\Delta}) = 0$
  As in Case 1 and 2, (A.42) can be satisfied for all $\mathrm{Tr}(\mathbf{\Delta}) = 0$ provided that $\lambda\,\mathrm{G}(\mathbf{D}_0) > 0$. As a result, $\lambda\,\mathrm{G}(\mathbf{D}_0) > 0$ only if $\mathrm{G}(\mathbf{D}_0) > 0$ (**C4**).

Combining all three cases, (A.42) can be satisfied given that **C1**, **C2**, **C3** and **C4** are satisfied. Similarly it can be shown that the conditions **C1**, **C2**, **C3** and **C4** are sufficient for (A.42) to be true which concludes the proof of the Theorem 1. ■

REFERENCES

[1] D. L. Donoho, "Compressed Sensing," *Information Theory, IEEE Transactions on*, 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," *Signal Processing, IEEE Transactions on*, vol. 60, pp. 4658 –4671, sept. 2012.

[3] B. C. Ng and C. M. S. See, "Sensor-array calibration using a maximum-likelihood approach," *Antennas and Propagation, IEEE Transactions on*, vol. 44, pp. 827 –835, jun 1996.

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

[5] E. C. Johnson and D. L. Jones, "Joint Recovery of Sparse Signals and Parameter Perturbations with Parameterized Measurement Models," in *Acoustics Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*.

[6] E. J. Candès, 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.

[7] M. Herman and T. Strohmer, "General deviants: An analysis of perturbations in compressed sensing," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, pp. 342 –349, april 2010.

[8] U. S. Kamilov, A. Bourquard, E. Bostan, and M. Unser, "Autocalibrated Signal Reconstruction from Linear Measurements Using Adaptive GAMP," in *Acoustics Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*.

[9] C. Schülke, F. Caltegirone, F. Krzakala, and L. Zdeborová, "Blind Calibration in Compressed Sensing Using Message Passing Algorithms," *arXiv preprint arXiv:1306.4355*, 2013.

[10] E. J. Candès, Y. C. Eldar, T. Strohmer, and V. Voroninski, "Phase Retrieval via Matrix Completion," *Imaging Sciences, SIAM Journal on (to appear)*, 2011.

[11] E. J. Candès, T. Strohmer, and V. Voroninski, "PhaseLift: Exact and Stable Signal Recovery from Magnitude Measurements via Convex Programming," *Communications on Pure and Applied Mathematics (to appear)*, 2011.

[12] R. Balan, P. Casazza, and D. Edidin, "On signal reconstruction without phase," *Applied and Computational Harmonic Analysis*, vol. 20, pp. 345–356, May 2006.

[13] R. Balan, B. G. Bodmann, P. G. Casazza, and D. Edidin, "Fast algorithms for signal reconstruction without phase," in *Wavelets XII, Proc. of SPIE*, vol. 6701, (San Diago, California, USA), pp. 67011L.1–67011L.9, August 2007.

[14] 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.

[15] H. Ohlsson, A. Y. Yang, R. Dong, and S. S. Sastry, "Compressive Phase Retrieval from Squared Output Measurements via Semidefinite Programming," *arXiv preprint arXiv:1111.6323*, 2011.

[16] H. Ohlsson, A. Y. Yang, R. Dong, and S. S. Sastry, "CPRL-An Extension of Compressive Sensing to the Phase Retrieval Problem," in *Neural Information Processing Systems (NIPS)*, 2012.

[17] H. Ohlsson, A. Yang, R. Dong, and S. Sastry, "Quadratic Basis Pursuit," in *Signal Processing with Adaptive Sparse Structured Representations (SPARS) Workshop*, (Lausanne, Switzerland), 2013.

[18] X. Li and V. Voroninski, "Sparse Signal Recovery from Quadratic Measurements via Convex Programming," *arXiv preprint arXiv:1209.4785*, no. September, pp. 1–15, 2012.

[19] R. Gribonval, G. Chardon, and L. Daudet, "Blind Calibration for Compressed Sensing by Convex Optimization," in *Acoustics Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pp. 2713–2716, 2012.

[20] D. L. Donoho and J. 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, pp. 4273–93, Nov. 2009.

[21] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. 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.

[22] P. Schniter and S. Rangan, "Compressive Phase Retrieval via Generalized Approximate Message Passing," in *Proceedings of Allerton Conference on Communication, Control, and Computing*, 2012.

[23] S. Oymak, A. Jalali, M. Fazel, Y. C. Eldar, and B. Hassibi, "Simultaneously Structured Models with Application to Sparse and Low-rank Matrices," *arXiv preprint arXiv:1212.3753*, 2012.