COMPRESSIVE SIGNAL RECOVERY UNDER SENSING MATRIX ERRORS COMBINED WITH UNKNOWN MEASUREMENT GAINS

Jian Vora[1], Ajit Rajwade[2]

Department of Electrical Engg., IIT Bombay[1]
Department of Computer Science and Engg., IIT Bombay[2]

ABSTRACT

Compressed Sensing assumes a linear model for acquiring signals however imperfections may arise in the specification of the 'ideal' measurement model. We present the first study which considers the case of two such common calibration issues, when they occur *jointly*: (a) unknown measurement scaling (sensor gains) due to hardware vagaries or due to unknown object motion in MRI scanning, *in conjunction with* (b) unknown offsets to measurement frequencies in case of a Fourier measurement matrix. A Fourier measurement matrix measures the Fourier transform of the signal at pre-specified M base frequencies $\{u_i\}_{i=1}^{M}$. However, due to hardware issues, the system may measure the Fourier transform at frequencies $\{u_i+\delta_i\}_{i=1}^{M}$ where the offsets $\{\delta_i\}_{i=1}^{M}$ are unknown. We propose an alternating minimisation algorithm for on-the-fly signal recovery in the case when errors (a) and (b) occur *jointly*. We show simulation results over a variety of situations which outperform the baselines of signal recovery by ignoring either or both types of calibration errors. We also show theoretical results for signal recovery by introducing a perturbed version of the well-known Generalized Multiple Measurement Vectors (GMMV) model.

***Index Terms*—** Fourier Compressed Sensing, Sensor gains, Frequency perturbations, blind calibration

1. INTRODUCTION

Compressed Sensing (CS) has gained immense popularity for speedy acquisition of sparse signals and enjoys theoretical guarantees for signal recovery from compressive measurements [2]. Consider the model $\boldsymbol{y} = \boldsymbol{\Phi x}+\boldsymbol{\eta}$ where $\boldsymbol{y} \in \mathbb{C}^M$ is the measurement vector for a signal $\boldsymbol{x} \in \mathbb{C}^N$ and $\boldsymbol{\eta} \in \mathbb{C}^M$ is a noise vector. The measurements are acquired through a linear model $\boldsymbol{\Phi} \in \mathbb{C}^{M\times N}$ with $M < N$. It is assumed that the signal $\boldsymbol{x}$ is represented in some orthonormal basis $\boldsymbol{\Psi} \in \mathbb{C}^{N\times N}$, such that $\boldsymbol{x} = \boldsymbol{\Psi\theta}$ where $\boldsymbol{\theta}$ is a sparse vector. Moreover consider that $\boldsymbol{\Phi\Psi}$ follows the Restricted isometry property (RIP) due to which no sparse vector (other than the $\mathbf{0}$ vector) lies in the null-space of $\boldsymbol{\Phi\Psi}$ [2]. These two conditions guarantee the robust and stable recovery of signal $\boldsymbol{x}$ from $\boldsymbol{y}, \boldsymbol{\Phi}$ using convex programs such as Basis Pursuit Denoising (BPDN) which minimizes $\|\boldsymbol{\theta}\|_1$ such that $\|\boldsymbol{y}-\boldsymbol{\Phi\Psi\theta}\|_2 \leq \epsilon$ where ϵ is an upper bound on the noise level [2]. In this paper, we consider the case where $\boldsymbol{\Phi}$ is a partial Fourier matrix, henceforth denoted as $\boldsymbol{F}$. Most CS algorithms assume that the sensing matrix $\boldsymbol{F}$ is known correctly. However, it is very common to have calibration errors in the specifications of the forward model.

The first kind of errors we consider are **unknown multiplicative gains** in $\boldsymbol{y}$ due to the transfer functions of the sensing elements. This leads to the following model ($\forall i \in \{1, ..., M\}$):

$$y_i = g_i \sum_l e^{-\mathrm{j}2\pi \mathrm{u_i l}} x(l) + \eta_i \implies \boldsymbol{y} = \mathrm{diag}(\boldsymbol{g})\boldsymbol{F\Psi\theta} + \boldsymbol{\eta}, \tag{1}$$

where $\boldsymbol{g} \in \mathbb{C}^{\boldsymbol{M}}$ are the unknown gains in $\boldsymbol{y}$. This model is useful for a number of situations including gain calibration [3, 4], blind deconvolution [4], or object motion during MRI scanning [5]. In the last case, we have a Fourier measurement matrices. The measurement of a 1D signal displaced by d units is expressed as $y_i = \sum_l e^{-\mathrm{j}2\pi \mathrm{u_i l/N}} x(l + d) = g_i \sum_{\tilde{l}} e^{-\mathrm{j}2\pi \mathrm{u_i} \tilde{\mathrm{l}}/\mathrm{N}} x(\tilde{l})$, where $g_i \triangleq e^{\mathrm{j}2\pi \mathrm{u_i d/N}}, \mathrm{j} \triangleq \sqrt{-1}$ and $l, \tilde{l}$ are time indices. This leads to the same model as Eqn. 1.

The second error that we consider is that of **frequency perturbations** in a Fourier sensing matrix (common in MRI). However, the frequencies at which the system measures the Fourier transform may undergo unknown additive perturbations. (Eg: in case of MRI, these may be due to gradient delays [6]). This leads to the following model ($\forall i \in \{1, ..., M\}$):

$$y_i = \sum_l e^{-\mathrm{j}2\pi(\mathrm{u_i}+\delta_\mathrm{i})\mathrm{l/N}} x(l) + \eta_i \implies \boldsymbol{y} = \boldsymbol{F}(\boldsymbol{\delta})\boldsymbol{\Psi\theta} + \boldsymbol{\eta}, \tag{2}$$

where $\boldsymbol{F}(\boldsymbol{\delta})$ is the perturbed Fourier matrix and $\boldsymbol{\delta}$ is the vector of frequency perturbations.

The contribution of this paper is to consider the case when perturbations of the type in Eqns. 1 and 2 occur *simultaneously*, which occurs in MRI acquisition. The existing literature has treated these two problems separately which is outlined subsequently.

Email addresses of the authors are `jianvora@iitb.ac.in` and `ajitvr@cse.iitb.ac.in`. Code for replication of the results at [1]. AR thanks SERB Grant #10013890.

For example, simultaneous signal estimation and gain calibration methods have been developed in [4] via linear least squares, in [3] via the 'lifting technique', in [7] via many different convex estimators, and in [8] using non-convex optimization. The latter two techniques consider $q \geq 1$ different sets of measurements ('snapshots') for the estimation.

Our group's previous work in [9] considers Fourier CS with perturbed frequencies. A large body of work in CS with model perturbations considers the model $\boldsymbol{y} = (\boldsymbol{A}+\boldsymbol{B\Delta})\boldsymbol{x}+\boldsymbol{\eta}$ where $\boldsymbol{A}, \boldsymbol{B}$ are known and $\boldsymbol{\Delta}, \boldsymbol{x}$ are to be recovered [10, 11, 12, 13]. However as argued in detail in [9, Sec. 4], this problem is much more similar to one where the representation matrix $\boldsymbol{\Psi}$ and not the sensing matrix $\boldsymbol{\Phi}$ has been perturbed. An approach for sensing matrix perturbation via total least squares has been presented in [14]. However none of these approaches consider the problem *in conjunction with* unknown gains. The approach closest to the one we will present is [15] which considers unknown sensor gains and perturbations in the representation matrix $\boldsymbol{\Psi}$, and that too for the case of *multiple* measurement snapshots. They propose a non-convex estimator and prove uniqueness of the solutions in the case of infinitely many signal snapshots.

Contributions: To the best of our knowledge, this is the first work to consider both unknown gains and sensing matrix perturbations *jointly*. We propose a monolithic alternating minimization algorithm for such joint blind calibration. We also provide theoretical results for signal recovery in a generalized multiple measurement vector (GMMV) framework [16]. Our work here significantly extends our group's previous work on Fourier matrix perturbation from [9] in two ways: (1) we consider unknown gains, and (2) we derive theoretical bounds for arbitrary orthonormal $\boldsymbol{\Psi}$ whereas [9] presents theoretical results only for $\boldsymbol{\Psi} =$ identity matrix $\boldsymbol{I}$.

2. THEORY

We describe our perturbation model along with all underlying assumptions in Sec. 2.1, followed by presenting the estimator and the recovery algorithm in Sec. 2.2.

2.1. MODEL FOR SIGNAL ACQUISITION UNDER PERTURBATIONS

Combining the models from Eqns. 1 and 2, our final sensing model can be described as follows:

$$\boldsymbol{y} = \text{diag}(\boldsymbol{g})\boldsymbol{F}(\delta)\boldsymbol{\Psi}\boldsymbol{\theta} + \boldsymbol{\eta}. \tag{3}$$

Thus given just a *single* measurement vector $\boldsymbol{y}$, we need to recover the gains in $\boldsymbol{g}$, the frequency perturbations in $\boldsymbol{\delta}$ and the signal coefficients $\boldsymbol{\theta}$ simultaneously. In many practical applications, each perturbation δ_i is not independent and often $\boldsymbol{\delta}$ is governed by some $P \ll M$ 'perturbation parameters'(PPs), following [9]. The PPs are denoted by $\boldsymbol{\beta} \triangleq \{\beta_k\}_{k=1}^{P}$. Thus we have $\forall i \in \{1, ..., n\} \exists k \in \{1, ..., P\}$ s. t. $\delta_i = h_p(u_i, \beta_k)$ where the function h_p depends on the particular application. For our recovery algorithm, we assume that $\forall k, 1 \leq k \leq P, |\beta_k| \leq r_\beta$ for some $r_\beta \in \mathbb{R}_{\geq 0}$. In the same way as frequency perturbations, we consider two cases for the number of independent gains: (1) All M gain values are free parameters due to arbitrary sensor gains, and (2) The gain values are governed by $Q \ll M$ shared 'gain parameters' $\{\gamma_k\}_{k=1}^{Q}$ (GPs) so that $\forall i \in \{1, ..., n\} \exists k \in \{1, ..., Q\}$ s. t. $g_i = h_g(u_i, \gamma_k)$ where the function h_g is application-dependent. For example, in the case of motion in MRI as described earlier, $Q = 1$ as the single GP d governs the various gain perturbations. Furthermore, we assume that $\forall k, 1 \leq k \leq M, g_{min} < |g_k| < g_{max}$, and also that $g_k \neq 0$ since a measurement with $g_k = 0$ contains no information about the signal and can be discarded.

2.2. Recovery Algorithm

We propose to use an SQ-LASSO-based estimator [17] for estimating the signal and perturbations simultaneously. Our estimator minimizes the following cost function:

$$J(\boldsymbol{\theta}, \boldsymbol{\delta}, \boldsymbol{g}) := \|\boldsymbol{\theta}\|_1 + \lambda\|\boldsymbol{y} - \text{diag}(\boldsymbol{g})\boldsymbol{F}(\boldsymbol{\delta})\boldsymbol{\Psi}\boldsymbol{\theta}\|_2,$$
$$\text{where } \boldsymbol{\delta} \in [-r, r]^M, \boldsymbol{g} \in [g_{min}, g_{max}]^M. \tag{4}$$

We have used SQ-LASSO instead of the LASSO because the former allows tuning of the regularization parameter λ in a manner independent of the noise level [17]. We propose an easy-to-implement three-way alternating minimization procedure to minimize the cost function in Eqn. 4 as mentioned in Algorithm 1. Each of the three major steps in Alg. 1 can be executed efficiently. The estimation of $\boldsymbol{\theta}$ given fixed $\boldsymbol{g}, \boldsymbol{\delta}$ is a convex problem and can be implemented via the CVX library. The estimation of $\boldsymbol{\delta}$ and $\boldsymbol{g}$ can be performed by brute-force search in their respective ranges. Note that these values can be searched *independently* for different measurements (or groups of measurements if we consider the case of PPs in $\boldsymbol{\beta}$). As this is a non-convex optimization problem, the procedure could get stuck in local optima. To mitigate this problem, we employ a multi-start strategy with different initial conditions and choose the solution $(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\delta}}, \hat{\boldsymbol{g}})$ which minimizes $J(\cdot)$.

3. EMPIRICAL EVALUATION

In this section, we present simulation results for signal recovery using Alg. 1. Note that in the presence of unknown gains in $\boldsymbol{g}$, it is possible to obtain the solution only up to a scalar multiple, i.e, if $(x, \boldsymbol{g}, \boldsymbol{\delta})$ yields a measurement vector $\boldsymbol{y}$, then $(x/\alpha, \alpha\boldsymbol{g}, \boldsymbol{\delta})$ also yields the same measurement vector for any $\alpha \in \mathbb{C}$ and the two solutions are indistinguishable. If $\hat{\boldsymbol{x}}$ is the estimate of the true signal $\boldsymbol{x}$, we quantify the performance of the algorithm by computing the normalized error value NE $\triangleq \|\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2} - \frac{\hat{\boldsymbol{x}}}{\|\hat{\boldsymbol{x}}\|_2}\|_2$ to overcome the scale ambiguity. For simplicity, we consider that

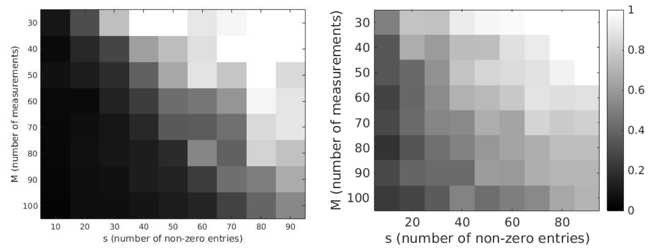
$h_p(u_i, \beta_k) = \beta_k, h_g(u_i, \gamma_k) = \gamma_k$ as well as real-valued gains randomly sampled from Uniform$(1-r_g, 1+r_g)$ where $0 < r_g \leq 1$. In all experiments, the regularization parameter λ was selected using cross validation [18]. Also, iid noise (in $\boldsymbol{\eta}$) generated from $\mathcal{N}(0, f_n^2 y_{n,avg}^2)$ was added to the CS measurements, where $0 < f_n < 1$ and $y_{n,avg}$ is the average absolute value of the noiseless measurements.

We first consider an experiment to recover a sparse signal $\boldsymbol{x}$ with $n = 100$ elements and non-zero elements drawn from $\mathcal{N}(0, 25)$ at randomly chosen indices, for many different values of M and $\|\boldsymbol{x}\|_0$. The results for this experiment are plotted in Fig. 1, with NE values averaged over 5 signals. Each grid entry in Fig. 1 shows the NE value for a chosen $(M, \|\boldsymbol{x}\|_0)$, and the trends clearly show reduction in NE when M increases and $\|\boldsymbol{x}\|_0$ decreases. For every configuration, we chose the number of PPs and GPs to be $P = Q = 5$ with $r_\beta = 1, r_g = 0.4, f_n = 0.05$.

In a second experiment, we consider a signal of length $n = 200$ with compressive measurements obtained with a different GP per measurement (i.e., $Q = M$) and $P = 15, r_\beta = 0.5, r_g = 0.4, f_n = 0.05$. The NE results for this experiment are plotted in Fig. 2. Owing to larger Q, the NE values in Fig. 2 are larger than those in Fig. 1.

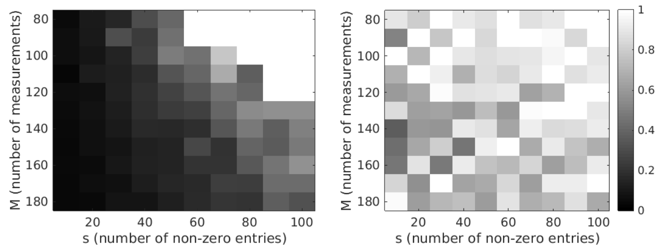
Algorithm 1 Simultaneous recovery of signal, gain and frequency perturbation parameters

procedure ALTERNATINGRECOVERY
  $converged \leftarrow False$
  $\boldsymbol{\delta} \leftarrow$ random sample from Unif$[-r_{max}, +r_{max}]$
  $\boldsymbol{g} \leftarrow$ random sample from Unif$[g_{min}, g_{max}]$
  while $converged == False$ **do**
    $\hat{F} \leftarrow F_{\boldsymbol{\delta}}$
    $\hat{D} \leftarrow \text{diag}(\boldsymbol{g})$
    Estimate $\boldsymbol{\theta}$ as: $\min_{\boldsymbol{\theta}} \|\boldsymbol{\theta}\|_1 + \lambda\|y - \hat{D}\hat{F}\Psi\boldsymbol{\theta}\|_2$
    for k in $1 \rightarrow T$ **do**
      Test each discretized value of δ_k in range $-r_{max}$ to r_{max} and select the value to achieve the following optimum
      $$\min_{\delta_k} \|y_k - \hat{D}\hat{F}(\delta_k)\Psi\theta\|_2$$
      Update $\hat{\delta_k} \leftarrow \delta_k$
    for k in $1 \rightarrow T$ **do**
      Test each discretized value of g_k in range g_{min} to g_{max} and select value to achieve the following optimum
      $$\min_{g_k} \|y_k - \hat{D}(g_k)\hat{F}\Psi\alpha\|_2$$
      Update $\hat{g_k} \leftarrow g_k$
    if $\|\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}_{prev}\|_2 < 1\epsilon^{-2}$ **and** $\|\hat{\boldsymbol{\delta}} - \hat{\boldsymbol{\delta}}_{prev}\|_2 < 1\epsilon^{-2}$ **and** $\|\hat{\boldsymbol{g}} - \hat{\boldsymbol{g}}_{prev}\|_2 < 1\epsilon^{-2}$ **then**
      $converged \leftarrow True$
  return $\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\delta}}, \hat{\boldsymbol{g}}$



(a) (b)

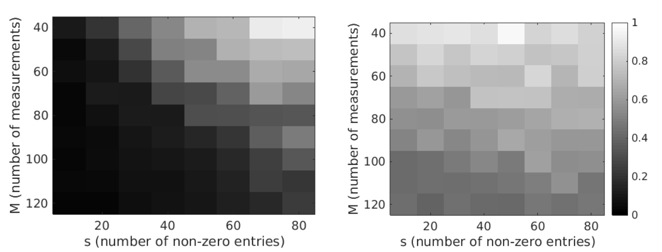
Fig. 1. NE values vs. $(M, \|\boldsymbol{x}\|_0)$ for CS recon. with $P = Q = 5$ using (a) Alg. 1 and (b) Baseline uncalibrated LASSO.



(a) (b)

Fig. 2. NE values vs. $(M, \|\boldsymbol{x}\|_0)$ for CS recon. with $P = 15, Q = M$ using (a) Alg. 1 and (b) Baseline uncalibrated LASSO.

Finally, we consider the important case when the signal is sparse in some other orthonormal basis (i.e., $\boldsymbol{\Psi} \neq \boldsymbol{I}$) like the Haar wavelet. Our reconstruction results for this are shown on signals with $n = 128$ elements. Other parameter values were $P = 5, Q = M, r_\beta = 0.1, r_g = 0.2, f_n = 0.02$. The results for this experiment are shown in Fig. 3.



(a) (b)

Fig. 3. NE values vs. $(M, \|\boldsymbol{x}\|_0)$ for CS recon. with $P = 5, Q = M$ using (a) Alg. 1 and (b) Baseline uncalibrated LASSO. The signal is sparse in Haar wavelet basis.

It is clear from Figs. 1, 2, 3, that our algorithm significantly outperforms the obvious baseline which ignored the calibration errors. Further empirical results can be found in the supplemental material [19] in which we compare the results of our recovery algorithm with two different intermediate baselines (i) only calibrating the δs and (ii) only calibrating the gains. These are implemented using Alg. 1 but without minimising over the gains or δs respectively.

4. THEORETICAL GUARANTEES

In this section, we present some theoretical guarantees for signal reconstruction in the case of the model in Eqn. 3 but with noiseless measurements (i.e. $\boldsymbol{\eta} = \mathbf{0}$). As argued in Sec. 3, we can expect signal reconstruction only up to a scalar multiplication. We term the solution 'unique' if we can obtain the signal up to this inevitable scale ambiguity. We shall first show that such uniqueness holds for sparsity in the canonical basis. We then present performance bounds if signal $\boldsymbol{x}$ is sparse in an arbitrary orthonormal basis $\boldsymbol{\Psi}$.

Canonical Basis: As per the model in Eqn. 3, we have:

$$y_i = \sum_l (g_i x(l) e^{-\mathrm{j}2\pi\delta_\mathrm{i}\mathrm{l}/\mathrm{N}}) e^{-\mathrm{j}2\pi\mathrm{u_i}\mathrm{l}/\mathrm{N}}. \tag{5}$$

As per the assumptions in Sec. 2.1, the (g_i, δ_i) values may not be different for each of the M measurements, in many applications. Let us consider that there are only $T \ll M$ unique values for (g_i, δ_i), and divide $\boldsymbol{y}$ into T disjoint sub-vectors $\{\boldsymbol{y^{(k)}}\}_{k=1}^T$, each consisting of M/T measurements all corresponding to the same pair of parameter values from $(\boldsymbol{g}, \boldsymbol{\delta})$. The corresponding measurement sub-matrices are denoted as $\{\boldsymbol{F^{(k)}}\}_{k=1}^T$. The forward model for $\boldsymbol{y^{(k)}}$ is:

$$\boldsymbol{y^{(k)}} = \boldsymbol{F^{(k)}} \boldsymbol{z_k}, \tag{6}$$

where we define signal $\boldsymbol{z_k}$ such that $z_{k,l} \triangleq g_k x(l) e^{-\mathrm{j}2\pi\delta_\mathrm{k}\mathrm{l}/\mathrm{N}}$. Since different signals $\{\boldsymbol{z_k}\}_{k=1}^T$ have the same support, their estimation from $(\{\boldsymbol{y^{(k)}}, \boldsymbol{F^{(k)}}\}_{k=1}^T)$ is a multiple measurement vector (MMV) problem [20, Sec. V-A]. However the measurement matrices $\{\boldsymbol{F^{(k)}}\}_{k=1}^T$ are different, and hence this is a *generalized* MMV (GMMV) problem [16, 21]. If M/T is large enough, we can recover the signals $\{\boldsymbol{z_k}\}_{k=1}^T$ faithfully as per [21, Thm. 23, Chapter 5].

But our final aim is to recover the signal $\boldsymbol{x}$ uniquely, from the intermediate signals $\{\boldsymbol{z_k}\}_{k=1}^T$. Let us express the unknown gain g_k associated with $\boldsymbol{z_k}$ as $g_k = a_k e^{jb_k}$ where a_k, b_k are the magnitude, phase respectively. To argue for uniqueness, we proceed by contradiction. Let us suppose that given $\boldsymbol{z_k}$, we have two estimates of the signal $\boldsymbol{x}$, the PP δ_k and the GP g_k, denoted as $(\boldsymbol{x_1}, \delta_{k1}, g_{k1} = a_{k1} e^{jb_{k1}})$ and $(\boldsymbol{x_2}, \delta_{k2}, g_{k2} = a_{k2} e^{jb_{k2}})$. Thus we want to examine the conditions under which, we would obtain:

$$a_{k1} x_1(l) e^{j(b_{k1} - 2\pi\delta_{k1} l/N)} = a_{k2} x_2(l) e^{j(b_{k2} - 2\pi\delta_{k2} l/N)} \tag{7}$$

$$\implies \frac{a_{k1} x_1(l)}{a_{k2} x_2(l)} = e^{-j[2\pi(\delta_{k2} - \delta_{k1}) l/N + (b_{k1} - b_{k2})]}. \tag{8}$$

If the signal $\boldsymbol{x}$ is real-valued, then the LHS is a real quantity so we want $\forall l, 2\pi(\delta_{k2} - \delta_{k1}) l/N + (b_{k1} - b_{k2}) = c_1 \pi$ for all values of l and for some integer c_1. For $l = 0$, we obtain $b_{k1} - b_{k2} = c_1 \pi$. Hence, for other values of l we obtain $2\pi(\delta_{k2} - \delta_{k1}) l/N = c_2 \pi$ for some integer c_2. Following further arguments from Lemma 1 of [9], we can recover the signal $\boldsymbol{x}$. We thus arrive at the following Lemma for recovery of $\boldsymbol{x}$ from measurements in Eqn. 1.

Lemma 1: Consider measurements of the form $\boldsymbol{y} = g\boldsymbol{F}(\delta)\boldsymbol{x}$, where (a) $\boldsymbol{x} \in \mathbb{R}^N$ is an s-sparse vector (b) δ, g are a single scalar PP and GP respectively. Then $\boldsymbol{x}$ can be recovered up to a scalar multiple if $\boldsymbol{y}$ has at least $\mathcal{O}(s\log^2 s\log N)$ measurements and $|\delta| < N/4$. If $\boldsymbol{x} \in \mathbb{C}^N$, then we can only recover the modulus of each element of the signal upto a scalar multiple.■

Arbitrary Orthonormal Basis: In this case, we first define the $N \times N$ matrix $\boldsymbol{\Delta_k} \triangleq \mathrm{diag}(e^{-\mathrm{j}2\pi\delta_\mathrm{k}\mathrm{l}/\mathrm{N}})$. We present our forward model as follows, for every $1 \leq k \leq T$:

$$\begin{aligned} \boldsymbol{y^{(k)}} &= g_k \boldsymbol{F^{(k)}} \boldsymbol{\Delta_k} \boldsymbol{\Psi} \boldsymbol{\theta} \\ &= (g_k \boldsymbol{F^{(k)}} \boldsymbol{\Psi} + g_k \boldsymbol{F^{(k)}} \boldsymbol{\Psi} (\boldsymbol{\Psi}^H \boldsymbol{\Delta_k} \boldsymbol{\Psi} \boldsymbol{\Delta_k}^H - \boldsymbol{I})) \boldsymbol{\Delta_k} \boldsymbol{\theta} \\ &= \boldsymbol{A_k} \boldsymbol{z_k} + \boldsymbol{E_k} \boldsymbol{z_k}, \end{aligned} \tag{9}$$

where we define $\boldsymbol{A_k} \triangleq \boldsymbol{F^{(k)}} \boldsymbol{\Psi}$, $\boldsymbol{E_k} \triangleq \boldsymbol{F^{(k)}} \boldsymbol{\Psi} (\boldsymbol{\Psi}^H \boldsymbol{\Delta_k} \boldsymbol{\Psi} \boldsymbol{\Delta_k}^H - \boldsymbol{I}))$, and as defined earlier, we use $\boldsymbol{z_k} \triangleq g_k \boldsymbol{\Delta_k} \boldsymbol{\theta}$. Here, we regard $\boldsymbol{E_k}$ to be the additional perturbation matrix which will be a $\mathbf{0}$ matrix when the system is perfectly calibrated. Eqn. 9 presents a GMMV problem just like Eqn. 6, but further introduces signal-dependent noise $\boldsymbol{E_k} \boldsymbol{z_k}$. Due to this noise term, we are unable to present uniqueness results. However we can present performance bounds on $\boldsymbol{z_k}$ by invoking Theorem 24 of [21].

For such a performance bound to be useful, we would like the ratio $\epsilon_1 \triangleq \max_k \frac{\|\boldsymbol{E_k}\|_2}{\|\boldsymbol{A_k}\|_2}$ (the level of the signal-dependent noise) to not be too large, following [22]. We can show that $\epsilon_1 \leq \max_k \|\boldsymbol{\Psi}^H \boldsymbol{\Delta_k} \boldsymbol{\Psi} \boldsymbol{\Delta_k}^H - \boldsymbol{I}\|_2$ (details in suppl. material [19]). If we assume that $r \approx 0$, then we can employ a Taylor series approximation to study how ϵ_1 varies as a function of r (details in [19]). This is also illustrated in [19]. Moreover, consider that we have additive measurement noise $\boldsymbol{\eta}$ in $\boldsymbol{y}$ such that $\|\boldsymbol{\eta}\|_2 \leq \epsilon_2$. Then we have the upper bound on the noise in Eqn. 9 given by $\|\boldsymbol{E_k} \boldsymbol{z_k} + \boldsymbol{\eta_k}\|_2 \leq \epsilon \triangleq \epsilon_1 g_{max} \|\boldsymbol{x}\|_\infty + \epsilon_2$. We can use this upper bound for a BPDN algorithm in a GMMV framework as in [21, Thm. 24]. This yields recovery bounds for $\boldsymbol{z_k}$ and thereby for $\boldsymbol{x}$ via Lemma 1. As discussed earlier, if we can recover $\boldsymbol{z_k}$ from the above system of equations, then we can recover the signal $\boldsymbol{x}$.

5. CONCLUSION

Our paper is the first piece of work to have presented an algorithm, experimental results and theoretical results for the problem of compressive recovery in the face of two *combined* model-specification errors: unknown measurement gains and perturbations in the specified frequencies in a Fourier sensing matrix. Such combined errors arise in MRI acquisition with object motion. Possible avenues for future work could include (1) exploring an even more general perturbation model with errors in the representation matrix $\boldsymbol{\Psi}$ as well, or (2) strengthening the bounds using the additional structure in the signals for GMMV recovery apart from only common support.

6. REFERENCES

[1] J. Vora, "Code for implementation," `https://github.com/jianvora/CS-Calibrations`.

[2] E. Candès and M. Wakin, "An introduction to compressive sampling," *IEEE signal processing magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[3] S. Ling and T. Strohmer, "Self-calibration and biconvex compressive sensing," *Inverse Problems*, vol. 31, 2015.

[4] S. Ling and T. Strohmer, "Self-calibration via linear least squares," *CoRR*, vol. abs/1611.04196, 2016.

[5] Z. Yang, C. Zhang, and L. Xie, "Sparse MRI for motion correction," in *International Symposium on Biomedical Imaging*, 2013, pp. 962–965.

[6] E. Brodsky, A. Samsonov, and W. Block, "Characterizing and correcting gradient errors in non-cartesian imaging: Are gradient errors linear-time-invariant?," *Magnetic Resonance Imaging*, vol. 62, pp. 1466–1476, 2009.

[7] C. Bilen, G. Puy, R. Gribonval, and L. Daudet, "Convex optimization approaches for blind sensor calibration using sparsity," *IEEE Transactions on Signal Processing*, vol. 62, no. 8, pp. 4847–4856, 2014.

[8] V. Cambareri and L. Jacques, "A non-convex blind calibration method for randomised sensing strategies," in *4th International Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing (CoSeRa)*, 2016, p. 16–20.

[9] H. Pandotra, E. Malhotra, A. Rajwade, and K. S. Gurumoorthy, "Dealing with frequency perturbations in compressive reconstructions with fourier sensing matrices," *Signal Processing*, vol. 165, pp. 57 – 71, 2019.

[10] Z. Tan, Y Peng, and A Nehorai, "Joint sparse recovery method for compressed sensing with structured dictionary mismatches," *IEEE Transactions on Signal Processing*, vol. 62, no. 19, pp. 4997–5008, 2014.

[11] A. Fannjiang and H.-C. Tseng, "Compressive radar with off-grid targets: a perturbation approach," *Inverse Problems*, vol. 29, no. 5, 2013.

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

[13] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Trans. Information Theory*, vol. 59, no. 11, pp. 7465–7490, 2013.

[14] H. Zhu, G. Leus, and G. Giannakis, "Sparsity-cognizant total least-squares for perturbed compressive sampling," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, 2011.

[15] Y. C. Eldar, W. Liao, and S. Tang, "Sensor calibration for off-the-grid spectral estimation," *Applied and Computational Harmonic Analysis*, vol. 48, no. 2, pp. 570–598.

[16] R. Heckel and H. Bölcskei, "Joint sparsity with different measurement matrices," in *Allerton Conference on Communication, Control, and Computing*, 2012, pp. 698–702.

[17] A. Belloni, V. Chernozhukov, and L. Wang, "Square-root LASSO: pivotal recovery of sparse signals via conic programming," *Biometrika*, vol. 98, no. 4, pp. 791, 2011.

[18] J. Zhang, L. Chen, P. T. Boufounos, and Y. Gu, "On the theoretical analysis of cross validation in compressive sensing," in *ICASSP*, 2014.

[19] J. Vora and A. Rajwade, "Supplemental material," `https://www.cse.iitb.ac.in/~ajitvr/ICASSP2021_supp/jian_supp.pdf`.

[20] M. F. Duarte and Y. C. Eldar, "Structured compressed sensing: From theory to applications," *Trans. Sig. Proc.*, vol. 59, no. 9, pp. 4053–4085, 2011.

[21] R. Heckel, *Sparse signal processing: subspace clustering and system identification*, Ph.D. thesis, ETH Zurich, 2014.

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