General Perturbations in Compressed Sensing

Matthew Herman and Thomas Strohmer
Department of Mathematics, University of California, Davis

Introduction

We analyze the Basis Pursuit recovery method when observing signals with general perturbations (i.e., ***additive***, as well as ***multiplicative noise***). This **completely perturbed model** extends the previous work of Candès, Romberg and Tao on stable signal recovery from incomplete and inaccurate measurements. Our results show that, under suitable conditions, the stability of the recovered signal is limited by the noise level in the observation. Moreover, this accuracy is within a constant multiple of the best-case reconstruction using the technique of least squares.

Previous Work on Perturbations and ℓ_1-Minimization

Employing the techniques of compressed sensing (CS) to recover signals with a sparse representation has enjoyed a great deal of attention over the last 5–10 years. The initial studies considered an ideal unperturbed scenario:

$$\boldsymbol{b} \;=\; \boldsymbol{A}\boldsymbol{x}.$$

Here $\boldsymbol{b} \in \mathbb{C}^m$ is the observation vector, $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ is a full-rank measurement matrix or system model (with $m \le n$), and $\boldsymbol{x} \in \mathbb{C}^n$ is the signal of interest which has a sparse, or almost sparse, representation under some fixed basis. More recently researchers have included an ***additive noise*** term $\boldsymbol{e}$ into the received signal [C, CRT] creating a **partially perturbed model**:

$$\hat{\boldsymbol{b}} \;=\; \boldsymbol{A}\boldsymbol{x} + \boldsymbol{e} \tag{1}$$

This type of noise generally models simple errors which are uncorrelated with the data.

Given perturbed observation $\hat{\boldsymbol{b}}$ we are concerned with solving the Basis Pursuit (BP) problem [CDS]:

$$\boldsymbol{z}^\star = \operatorname*{argmin}_{\hat{\boldsymbol{z}}} \|\hat{\boldsymbol{z}}\|_1 \quad \text{such that} \quad \|\boldsymbol{A}\hat{\boldsymbol{z}} - \hat{\boldsymbol{b}}\|_2 \le \varepsilon' \tag{2}$$

for some $\varepsilon' \ge 0$.

The ***restricted isometry property*** (RIP) [CRT] for any matrix $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ defines, for each integer $K = 1, 2, \ldots$, the ***restricted isometry constant*** (RIC) δ_K, which is the smallest nonnegative number such that

$$(1-\delta_K)\|\boldsymbol{x}\|_2^2 \;\le\; \|\boldsymbol{A}\boldsymbol{x}\|_2^2 \;\le\; (1+\delta_K)\|\boldsymbol{x}\|_2^2 \tag{3}$$

holds for any K-sparse vector $\boldsymbol{x}$.

Theorem 0 [C]: Assume $\delta_{2K} < \sqrt{2} - 1$ for matrix $\boldsymbol{A}$ and noise with bounded energy $\|\boldsymbol{e}\|_2 \le \varepsilon'$. Then given partially perturbed observation $\hat{\boldsymbol{b}}$ (1) the stability of the solution to the BP problem (2) obeys

$$\|\boldsymbol{z}^\star - \boldsymbol{x}\|_2 \;\le\; C_0 K^{-1/2}\|\boldsymbol{x} - \boldsymbol{x}_K\|_1 + C_1\varepsilon'. \tag{4}$$

New Work: Incorporating Perturbations to $\boldsymbol{A}$

Now suppose that the sensing matrix $\boldsymbol{A}$ has also been perturbed:

$$\hat{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{E}.$$

- It is important to consider this kind of deviation since it can account for precision errors when applications call for physically implementing the matrix $\boldsymbol{A}$ in a sensor.
- For blind source separation $\boldsymbol{E}$ can absorb errors in estimating the mixing matrix $\boldsymbol{A}$.
- When $\boldsymbol{A}$ represents a system model, such as in the context of radar or telecommunications, then $\boldsymbol{E}$ can absorb errors in assumptions made about the transmission channel.
- Perturbation $\boldsymbol{E}$ can potentially also model quantization errors arising from the discretization of analog signals.

In general, these perturbations can be characterized as ***multiplicative noise*** and are **more difficult** to analyze than simple additive noise since they are **correlated** with the signal of interest; observation $\hat{\boldsymbol{b}}$ will contain an extra noise term $\boldsymbol{E}\boldsymbol{x}$.

New Results

Our **completely perturbed model** extends (1) by also incorporating a perturbed sensing matrix:

$$\hat{\boldsymbol{b}} \;=\; (\boldsymbol{A} + \boldsymbol{E})\boldsymbol{x} + \boldsymbol{e}. \tag{5}$$

Important Notation

We quantify the perturbations in (5) with the following *relative bounds*

$$\frac{\|\boldsymbol{E}\|_2}{\|\boldsymbol{A}\|_2} \le \varepsilon_{\boldsymbol{A}}, \qquad \frac{\|\boldsymbol{E}\|_2^{(K)}}{\|\boldsymbol{A}\|_2^{(K)}} \le \varepsilon_{\boldsymbol{A}}^{(K)}, \qquad \frac{\|\boldsymbol{e}\|_2}{\|\boldsymbol{b}\|_2} \le \varepsilon_{\boldsymbol{b}}, \tag{6}$$

where $\|\cdot\|_2$ is the spectral norm, and $\|\cdot\|_2^{(K)}$ is the largest spectral norm over all *K-column submatrices*. Also define

$$\kappa_{\boldsymbol{A}} = \frac{\sigma_{\max}(\boldsymbol{A})}{\sigma_{\min}(\boldsymbol{A})}, \qquad \kappa_{\boldsymbol{A}}^{(K)} = \frac{\sqrt{1+\delta_K}}{\sqrt{1-\delta_K}}$$

respectively be the typical *ℓ_2-condition number* of matrix $\boldsymbol{A}$, and its *restricted K-column condition number*. The second quantity can be thought of as the worst condition number over all K-column submatrices of $\boldsymbol{A}$. Finally, let vector $\boldsymbol{x}_K \in \mathbb{C}^n$ be the best K-term approximation to $\boldsymbol{x}$. Then the effect of the "tail" $\boldsymbol{x}_{K^c} = \boldsymbol{x} - \boldsymbol{x}_K$ of the data can be measured by the ratios

$$r_K = \frac{\|\boldsymbol{x}_{K^c}\|_2}{\|\boldsymbol{x}\|_2}, \qquad a_K = \frac{r_K\|\boldsymbol{A}\|_2}{\sqrt{(1-r_K^2)(1-\delta_K)}}.$$

Theorem 1 (RIP for $\hat{\boldsymbol{A}}$) [HS]: For any $K = 1, 2, \ldots$, fix the RIC δ_K associated with $\boldsymbol{A}$, and the relative perturbation $\varepsilon_{\boldsymbol{A}}^{(K)}$ associated with $\boldsymbol{E}$ in (6). Then the RIC

$$\hat{\delta}_K := \left(1+\delta_K\right)\left(1+\varepsilon_{\boldsymbol{A}}^{(K)}\right)^2 - 1 \tag{7}$$

for matrix $\hat{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{E}$ is the smallest nonnegative constant such that

$$(1-\hat{\delta}_K)\|\boldsymbol{x}\|_2^2 \;\le\; \|\hat{\boldsymbol{A}}\boldsymbol{x}\|_2^2 \;\le\; (1+\hat{\delta}_K)\|\boldsymbol{x}\|_2^2 \tag{8}$$

holds for any K-sparse vector $\boldsymbol{x}$.

Theorem 2 (Basis Pursuit Stability from Completely Perturbed Observation) [HS]: Fix the relative perturbations $\varepsilon_{\boldsymbol{A}}, \varepsilon_{\boldsymbol{A}}^{(K)}, \varepsilon_{\boldsymbol{A}}^{(2K)}$ and $\varepsilon_{\boldsymbol{b}}$ in (6). Assume the RIC for matrix $\boldsymbol{A}$ satisfies

$$\delta_{2K} \;<\; \frac{\sqrt{2}}{\left(1+\varepsilon_{\boldsymbol{A}}^{(2K)}\right)^2} - 1, \tag{9}$$

and that signal $\boldsymbol{x}$ satisfies $\frac{\|\boldsymbol{x}_{K^c}\|_2}{\|\boldsymbol{x}_K\|_2} < \frac{\sqrt{1-\delta_K}}{\|\boldsymbol{A}\|_2}$.

Set the total noise term

$$\varepsilon' := \left(r_K\kappa_{\boldsymbol{A}}\varepsilon_{\boldsymbol{A}} + \frac{\kappa_{\boldsymbol{A}}^{(K)}\varepsilon_{\boldsymbol{A}}^{(K)}}{1-a_K} + \varepsilon_{\boldsymbol{b}}\right)\|\boldsymbol{b}\|_2. \tag{10}$$

Then given completely perturbed observation $\hat{\boldsymbol{b}}$ (5) the stability of the solution to the BP problem (2) obeys

$$\|\boldsymbol{z}^\star - \boldsymbol{x}\|_2 \;\le\; C_0K^{-1/2}\|\boldsymbol{x} - \boldsymbol{x}_K\|_1 + C_1\varepsilon'. \tag{11}$$

Remark: Theorem 2 generalizes Theorem 0. Indeed, if matrix $\boldsymbol{A}$ is unperturbed, then $\boldsymbol{E} = \boldsymbol{0}$ and $\varepsilon_{\boldsymbol{A}} = \varepsilon_{\boldsymbol{A}}^{(K)} = 0$. It follows that $\hat{\delta}_K = \delta_K$ in (7), and the RIPs for $\boldsymbol{A}$ in (3) and $\hat{\boldsymbol{A}}$ in (8) coincide. Moreover, condition (9) in Theorem 2 reduces to $\delta_K < \sqrt{2}-1$, and the total noise term (10) collapses to $\|\boldsymbol{e}\|_2 \le \varepsilon_{\boldsymbol{b}}\|\boldsymbol{b}\|_2 =: \varepsilon'$ (see (6)); both of these are identical to the assumptions in Theorem 0.

Traditional Least Squares Recovery

Let the subset $T \subseteq \{1, \ldots, n\}$ have cardinality $|T| = K$, and note the following T-*restrictions*: $\boldsymbol{A}_T \in \mathbb{C}^{m\times K}$ denotes the submatrix consisting of the columns of $\boldsymbol{A}$ indexed by the elements of T, and similarly for $\boldsymbol{x}_T \in \mathbb{C}^K$. Suppose the "oracle" case where we already know the support T of $\boldsymbol{x}_K$. The Least Squares (LS) problem consists of solving:

$$\boldsymbol{z}_T^\# = \operatorname*{argmin}_{\hat{\boldsymbol{z}}_T} \|\boldsymbol{A}_T\hat{\boldsymbol{z}}_T - \hat{\boldsymbol{b}}\|_2. \tag{12}$$

Least Squares Stability from Completely Perturbed Observation [HS]:

Fix the relative perturbations $\varepsilon_{\boldsymbol{A}}^{(K)}, \varepsilon_{\boldsymbol{A}}^{(2K)}$ and $\varepsilon_{\boldsymbol{b}}$ in (6). Assume the RIC for matrix $\boldsymbol{A}$ satisfies

$$\delta_{2K} \;<\; \frac{\sqrt{2}}{\left(1+\varepsilon_{\boldsymbol{A}}^{(2K)}\right)^2} - 1,$$

and that the relative additive perturbation satisfies $\quad \varepsilon_{\boldsymbol{b}} < \frac{\sqrt{1-\delta_K}}{\sqrt{1+\delta_K}}$.

Set the total noise term $\quad \zeta' := \left(\frac{\kappa_{\boldsymbol{A}}^{(K)}\varepsilon_{\boldsymbol{A}}^{(K)}}{1-a_K} + \varepsilon_{\boldsymbol{b}}\right)\|\boldsymbol{b}\|_2$.

Then given completely perturbed observation $\hat{\boldsymbol{b}}$ (5) the stability of the solution to the LS problem (12) obeys

$$\|\boldsymbol{z}^\# - \boldsymbol{x}\|_2 \;\le\; \|\boldsymbol{x} - \boldsymbol{x}_K\|_2 + C_2\zeta'. \tag{13}$$

Comparison of Basis Pursuit with Least Squares Recovery

To make the comparison fair, we need to assume that $\boldsymbol{x}$ is strictly K-sparse. Then $\boldsymbol{x}_K = \boldsymbol{x}$, $r_K = a_K = 0$, and both solutions (11), (13) enjoy stability which is proportional to the same noise factor:

$$\|\boldsymbol{z}^\star - \boldsymbol{x}\|_2 \;\sim\; \|\boldsymbol{z}^\# - \boldsymbol{x}\|_2 \;\approx\; \left(\kappa_{\boldsymbol{A}}^{(K)}\varepsilon_{\boldsymbol{A}}^{(K)} + \varepsilon_{\boldsymbol{b}}\right)\|\boldsymbol{b}\|_2.$$

Numerical Simulations of Basis Pursuit

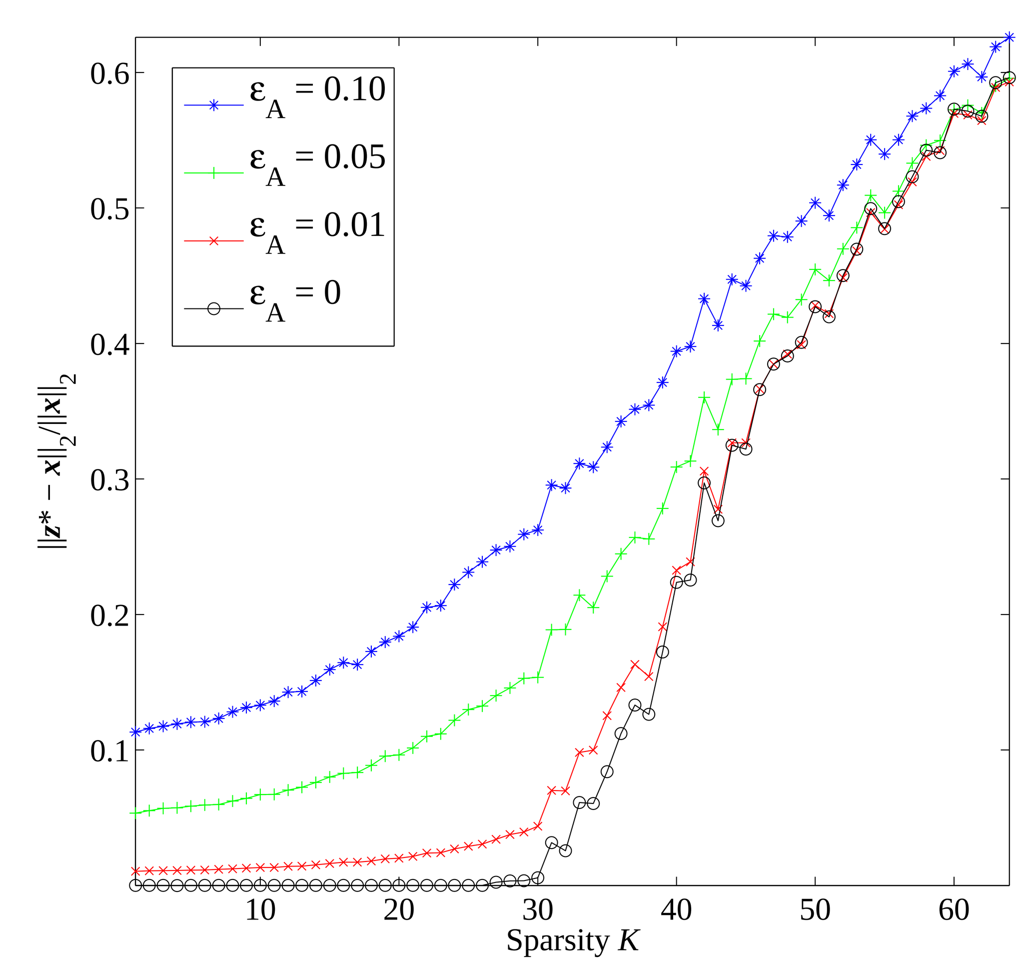

Figure 1: Average (100 trials) relative error of BP solution $\boldsymbol{z}^\star$ with respect to K-sparse $\boldsymbol{x}$ vs. Sparsity K for different perturbations $\varepsilon_{\boldsymbol{A}} = \|\boldsymbol{E}\|_2/\|\boldsymbol{A}\|_2$ where $\boldsymbol{E}, \boldsymbol{A} \in \mathbb{C}^{128\times 512}$ (and $\varepsilon_{\boldsymbol{b}} = 0$).

[CRT] E. J. Candès, J. Romberg, T. Tao Stable signal recovery from incomplete and inaccurate measurements. *Comm. Pure Appl. Math.*, 9, 1207–1223, 2006.

[C] E. J. Candès The Restricted Isometry Property and its Implications for Compressed Sensing. *Académie des Sciences*, I:346, 589-592, 2008.

[CDS] S. S. Chen, D. L. Donoho, M. A. Saunders Atomic Decomposition by Basis Pursuit. *SIAM Journal Sci. Comput.*, 20:1, 33–61, 1999.

[HS] M. Herman, T. Strohmer. General Deviants: An Analysis of Perturbations in Compressed. Submitted Feb. 2009.

Research supported by NSF Grant No. DMS-0811169 and NSF VIGRE Grant No. DMS-0636297.