ANALYZING CROSS VALIDATION IN COMPRESSED SENSING WITH MIXED GAUSSIAN AND IMPULSE MEASUREMENT NOISE WITH L1 ERRORS

Shubhang Bhatnagar [♯,*]*, Chinmay Gurjarpadhye* [♯,*] *and Ajit Rajwade*[†]

[*]Department of EE, [†]Department of CSE, IIT Bombay, India

ABSTRACT

Compressed sensing (CS) involves sampling signals at rates less than their Nyquist rates and attempting to reconstruct them after sample acquisition. Most such algorithms have parameters, for example the regularization parameter in LASSO, which need to be chosen carefully for optimal performance. These parameters can be chosen based on assumptions on the noise level or signal sparsity, but this knowledge may often be unavailable. In such cases, cross validation (CV) can be used to choose these parameters in a purely data-driven fashion. Previous work analysing the use of CV in CS has been based on the ℓ_2 cross-validation error with Gaussian measurement noise. But it is well known that the ℓ_2 error is not robust to impulse noise and provides a poor estimate of the recovery error, failing to choose the best parameter. Here we propose using the ℓ_1 CV error which provides substantial performance benefits given impulse measurement noise. Most importantly, we provide a detailed theoretical analysis and error bounds for the use of ℓ_1 CV error in CS reconstruction. We show that with high probability, choosing the parameter that yields the minimum ℓ_1 CV error is equivalent to choosing the minimum recovery error (which is not observable in practice). To our best knowledge, this is the first paper which theoretically analyzes ℓ_1-based CV in CS.

Index Terms— Compressed sensing, Cross validation, Impulse noise, L1 error

1. INTRODUCTION

The goal of compressed sensing (CS) is to improve the efficiency of signal acquisition by enabling a signal to be reconstructed from a small number of its samples [1]. It involves sampling the signal in a way such that most of the signal information is inherently available despite undersampling. The measured samples can be expressed in the form $\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x} + \boldsymbol{n}$, where $\boldsymbol{x} \in \mathbb{R}^N$ is a column vector representing the unknown signal, $\boldsymbol{y} \in \mathbb{R}^m$ is the vector of measurements, $\boldsymbol{n} \in \mathbb{R}^m$ is a noise vector such that $n_i \sim \mathcal{N}(0, \sigma_n^2/m)$, and $\boldsymbol{\Phi}$ is an $m \times N$ measurement matrix with $m \ll N$. The signal $\boldsymbol{x}$ is assumed to have a sparse representation in some $N \times N$ orthonormal basis $\boldsymbol{\Psi}$ so that $\boldsymbol{x} = \boldsymbol{\Psi}\boldsymbol{\theta}$ where $\boldsymbol{\theta} \in \mathbb{R}^N$ is a sparse vector.

The signal $\boldsymbol{x}$ can be reconstructed from its measurements $\boldsymbol{y}$ using a variety of techniques such as LASSO [2] or greedy algorithms like Orthogonal Matching pursuit (OMP) [3] and its variants. LASSO seeks to minimize the cost function $\|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x}\|_2^2 + \lambda\|\boldsymbol{x}\|_1$ w.r.t. $\boldsymbol{x}$, where λ is a regularization parameter. OMP seeks to minimize $\|\boldsymbol{x}\|_0$ s.t. $\|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x}\|_2 \leq \epsilon_n$ where ϵ_n is a regularization parameter dependent on the noise level. For optimal choice of the regularization parameter, many techniques rely on an estimate of the signal sparsity or noise level. However in real world scenarios, such information may not always be available. An alternative to this, proposed in [4], is using cross validation (CV), which is a purely data-driven technique. It proposes to set aside $m_{cv} < m$ measurements in $\boldsymbol{y}$ only for cross-validation (and not for reconstruction). These measurements are included in a $m_{cv} \times 1$ sub-vector $\boldsymbol{y_{cv}}$ with the corresponding $m_{cv} \times N$ sub-matrix $\boldsymbol{\Phi_{cv}}$. The remaining $m - m_{cv}$ measurements alone are used for compressive reconstruction, with each different value λ of the parameter chosen from a candidate set Λ. For each parameter value $\lambda \in \Lambda$, the CV error $\epsilon_{cv,\ell_2,\lambda} := \|\boldsymbol{y_{cv}} - \boldsymbol{\Phi_{cv}}\hat{\boldsymbol{x}}_{\boldsymbol{\lambda}}\|_2$ is computed where $\hat{\boldsymbol{x}}_{\boldsymbol{\lambda}}$ is an estimate of $\boldsymbol{x}$ using parameter λ in LASSO. The reconstruction $\hat{\boldsymbol{x}}_{\boldsymbol{\lambda}}$ corresponding to the value $\lambda \in \Lambda$ which yielded the lowest value of $\epsilon_{cv,\ell_2,\lambda}$ is chosen as the final one. The work in [5] proves theoretically using the Johnson-Lindenstrauss lemma [6] the close relationship between the data-driven CV error $\epsilon_{cv,\ell_2,\lambda}$ and the unobservable recovery error $\varepsilon_x := \|\boldsymbol{x} - \hat{\boldsymbol{x}}_{\boldsymbol{\lambda}}\|_2$ for the case of zero measurement noise. A detailed analysis for the case of additive iid Gaussian noise in the measurements $\boldsymbol{y}$ has been presented in [7], making use of the central limit theorem (CLT). The work in [7] calculates probabilistic bounds on the recovery error using the CV error, which theoretically justifies use of CV for CS in noisy signals.

In this paper, we analyse the use of CV for signals with mixed heavy-tailed and Gaussian noise in the measurement vector $\boldsymbol{y}$. The problem is motivated by the fact that heavy-tailed noise is common in many compressive systems (see [8, 9] and references therein). We observe that the ℓ_2-based CV error $\epsilon_{cv,\ell_2,\lambda}$ fails to give an accurate estimate of the actual recovery error for such noise models and fails to pick the best value of regularization parameters. We propose us-

♯ - CG and SB are both first authors with equal contribution. AR acknowledges support from SERB Grant #10013890.

ing the ℓ_1 based CV error $\epsilon_{cv,\ell_1,\lambda} := \|\boldsymbol{y_{cv}} - \boldsymbol{\Phi_{cv}}\hat{\boldsymbol{x}}_{\boldsymbol{\lambda}}\|_1$ instead, and demonstrate its significant performance benefits given impulse noise in the measurements. Most importantly, we derive the distribution of $\epsilon_{cv,\ell_1,\lambda}$ theoretically and establish its relationship to the recovery error $\|\boldsymbol{x} - \hat{\boldsymbol{x}}_{\boldsymbol{\lambda}}\|_2$. As will be seen, our theoretical results match numerical simulation results very closely.

2. PROBLEM FORMULATION AND MOTIVATION

The main aim of this section is to motivate the need for ℓ_1 CV errors. Throughout this work, we use a Gaussian sensing matrix, i.e $\forall(i,j), \Phi_{ij} \sim \mathcal{N}(0, 1/m)$, though the analysis can be easily extended to other sub-Gaussian sensing matrices. As mentioned earlier, we consider a noise model which consists of a mixture of additive iid Gaussian noise from $\mathcal{N}(0, \sigma_n^2/m)$ and impulse noise. The latter is modeled as the product of a discrete random variable which takes on values in $\{-1, +1, 0\}$ and a Gaussian random variable with a large mean as compared to $\sigma_n, \|\boldsymbol{x}\|_\infty$. Specifically, η_i which is the i^{th} element of the noise vector $\boldsymbol{\eta}$, is given as,

$$\eta_i = n_i + B_i G_i, \tag{1}$$

$$B_i = \begin{cases} +1 & \text{w.p. } b/2 \\ 0 & \text{w.p. } 1-b \\ -1 & \text{w.p. } b/2, \end{cases} \tag{2}$$

where $n_i \sim \mathcal{N}(0, \sigma_n^2/m), G_i \sim \mathcal{N}(\mu_g, \sigma_g^2)$. Here, we choose the probability b to be small, due to the sparse nature of impulse noise in many applications, and μ_g is chosen to be large compared to $\sigma_n, \sigma_g, \|\boldsymbol{x}\|_\infty$, since impulse noise often has very large magnitude as compared to the signal. Next, we present simulation results for CS based recovery of a signal with different amounts of impulse noise in the measurements. We perform the reconstruction using the following version of the LASSO with an ℓ_1-based data fidelity term:

$$\hat{\boldsymbol{x}}_{\boldsymbol{\lambda}} = \text{argmin}_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x}\|_1 + \lambda\|\boldsymbol{x}\|_1, \tag{3}$$

in which we select the regularisation parameter λ based on (1) minimum ℓ_1 CV error, (2) minimum ℓ_2 CV error, and (3) the minimum actual recovery error (implausible in real world, but useful for benchmarking). We use the parameters $\mu_g = 700, \sigma_n = 0.5, \sigma_g = 100, m = 420, N = 1200, m_{cv} = 20, s = \|\boldsymbol{x}\|_0 = 50$, for the experiment. In Fig. 1, we conduct a comparison between the performance of ℓ_1- and ℓ_2-based CV for reconstruction using Eqn. 3 with different amounts of impulse noise in the signal, by varying $b \in \{0.01, 0.02, ..., 0.09, 0.1\}$. In this experiment, the non-zero elements of $\boldsymbol{x}$ were chosen iid from $\mathcal{N}(0, 10)$. The results

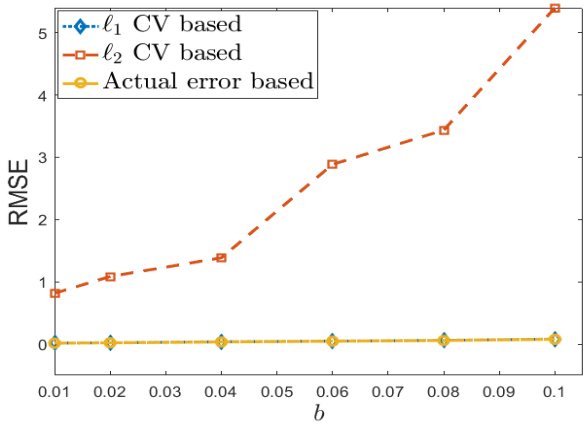

Fig. 1. Using ℓ_1-based CV error for parameter selection gives a much better RMSE than ℓ_2-based CV error, in presence of different occurrence probabilities (i.e., b) of impulse noise. The plots for ℓ_1-based CV error and true error nearly overlap.

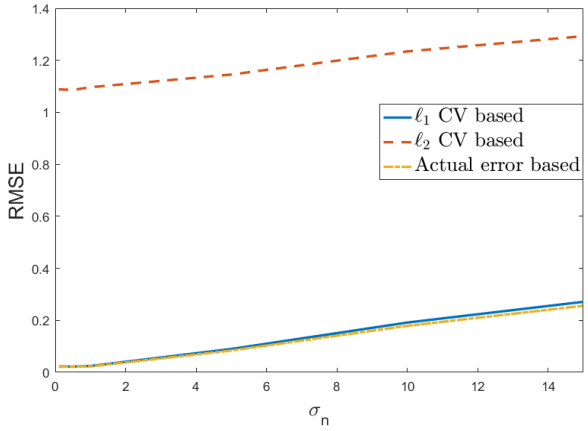

Fig. 2. Using ℓ_1-based CV error for parameter selection gives a much better RMSE than ℓ_2-based CV error, in presence of varying amounts of Gaussian noise (σ_n) given a fixed, small occurrence of impulse noise ($b = 0.02$).

in Fig. 1 show that the parameter λ chosen using ℓ_1-based CV error $\epsilon_{cv,\ell_1,\lambda}$ gives a reconstruction whose RMSE almost coincides with the actual error for a large number of possible values of b. This implies that ℓ_1-based CV error is fairly robust to impulse noise and provides a significantly superior estimate of the real recovery error compared to ℓ_2-based CV error.

Further, we compare the RMSE error for signal reconstruction using λ chosen using ℓ_1 based CV error and ℓ_2 based CV error for varying amount of non-impulse noise. For this experiment, we fix the probability of occurrence of impulse noise at $b = 0.02$ and vary σ_n from 0 to 15. All the other parameters are same as the last experiment and non-zero elements of $\boldsymbol{x}$ are chosen iid from $\mathcal{N}(0, 100)$. This figure further impresses the robustness of reconstruction using ℓ_1-based CV error compared to reconstruction using ℓ_2-based CV error.

3. THEORETICAL RESULTS

Having seen some experimental results which strongly motivate the use of ℓ_1-CV error, we now analytically derive a relationship between $\epsilon_{cv,\ell_1,\lambda} := \|\boldsymbol{y_{cv}} - \boldsymbol{\Phi_{cv}}\hat{\boldsymbol{x}}_{\boldsymbol{\lambda}}\|_1$ and the recovery error $\varepsilon_x := \|\boldsymbol{x} - \hat{\boldsymbol{x}}_{\boldsymbol{\lambda}}\|_2$. This relationship holds with high probability (via the CLT) for large values of m_{cv}. In Lemma 1, we derive the distribution of the $\epsilon_{cv,\ell_1,\lambda}$ in terms of ε_x and validate this result with relevant experiments. In Theorem 1, we use this distribution to obtain a high-probability confidence interval on the recovery error. We also give some experimental results to validate and illustrate the significance of these bounds. The distribution of the difference between the cross validation errors of two signal estimates is derived in Lemma 2 using CLT. This Lemma is crucially used to derive the probabilistic result in Theorem 2 which theoretically backs the use of ℓ_1-CV error to choose the optimal regularization parameter.

Lemma 1: Assuming that $\mu_g \gg \sigma_g, \sigma_n, \varepsilon_x$ and that m_{cv} is sufficiently large, we have $\epsilon_{cv,\ell_1,\lambda} \sim \mathcal{N}(\mu, \sigma^2)$, where

$\mu = bm_{cv}\mu_g + (1-b)m_{cv}\sqrt{\frac{2}{m\pi}(\varepsilon_x^2 + \sigma_n^2)}$,

$\sigma^2 = m_{cv}\Big(\frac{1}{m}(1-(1-b)^2\frac{2}{\pi})(\varepsilon_x^2+\sigma_n^2) + b(\sigma_g^2 + (1-b)\mu_g^2)$

$- 2b(1-b)\mu_g\sqrt{\frac{2}{m\pi}(\varepsilon_x^2+\sigma_n^2)}\Big)$. ■

The various parameters in Lemma 1 have been earlier defined in Sec. 2. The proof of this lemma (included in the suppl. material [10]) uses many properties of the absolute value of the Gaussian distribution [11, 12] followed by the CLT. Even if the CLT is an asymptotic result, we have observed excellent agreement between the empirically observed distribution of $\epsilon_{cv,\ell_1,\lambda}$ and the distribution predicted by Lemma 1, even at reasonable values of m_{cv}. This is seen in Fig. 3, where we empirically compute the pdf of $\epsilon_{cv,\ell_1,\lambda}$ for $N = 1200, m = 800, m_{cv} = 400$, given a single signal and 10^5 noise realizations. This clearly corroborates the correctness of Lemma 1. Also, we observe that our assumption that $\mu_g \gg \varepsilon_x$ typically holds for a wide range of values of the regularization parameter($\lambda \in \{0.001, 0.01, ...10000\}$) in our simulations. This is because impulses have very large magnitudes (as compared to the signal), and the reconstruction error is typically not of the same order as the impulse magnitude.

Theorem 1: Assuming that $\mu_g \gg \sigma_g, \sigma_n, \varepsilon_x$, m_{cv} is sufficiently large, and using ϵ_{cv} as shorthand for $\epsilon_{cv,\ell_1,\lambda}$, the following confidence interval holds with probability $\text{erf}(\varrho/\sqrt{2})$: $\frac{\sqrt{m}}{m_{cv}}\frac{\epsilon_{cv}-p(\varrho,+)}{h(\varrho,+)} \le \sqrt{\varepsilon_x^2+\sigma_n^2} \le \frac{\sqrt{m}}{m_{cv}}\frac{\epsilon_{cv}-p(\varrho,-)}{h(\varrho,-)}$., where

$$p(\varrho,\pm) := m_{cv}b\mu_G \pm \varrho\sqrt{m_{cv}b\,(\sigma_G^2 + (1-b)\mu_G^2)}$$

$$h(\varrho,\pm) := (1-b)\sqrt{\frac{2}{\pi}} \pm \varrho\sqrt{\frac{(1-(1-b)^2\frac{2}{\pi})}{m_{cv}}},$$

Here, $\text{erf}(u) := \frac{1}{\sqrt{\pi}}\int_{-u}^{u} \text{e}^{-t^2}dt$ denotes the error function and ϱ is a free parameter. One can observe that choosing a higher value of ϱ gives a looser bound but the bound holds with higher probability and vice versa for a lower value of ϱ. Furthermore, the width confidence interval from Theorem 1 denoted by $\mathcal{C}$ is given by $w := \frac{\sqrt{m}}{m_{cv}}\Big(\frac{\epsilon_{cv}(h(\varrho,+)-h(\varrho,-))}{h(\varrho,-)h(\varrho,+)}$ $+ \frac{p(\varrho,+)h(\varrho,-)-p(\varrho,-)h(\varrho,+)}{h(\varrho,-)h(\varrho,+)}\Big)$. As shown at the end of the proof of Theorem 1 in the supplemental material [10], this confidence interval drops to 0 in the limit as m_{cv} tends to infinity. The proof of Theorem 1 crucially uses Lemma 1 and is included in the supp. material [10]. In Fig. 4, we demonstrate the upper and lower bounds as per Theorem 1 and the empirical recovery error for $N = 1200, 40 \le m_{cv} \le 200, s = \|\boldsymbol{x}\|_0 = 50, b = 0.1, \sigma_n = 0.5, \mu_g = 700, \sigma_g = 100, \varrho = 3$. We note that similar results can be obtained for other parameters as well. For generating Fig. 4, we have averaged over 1000 instances, with new realizations of all random vari-

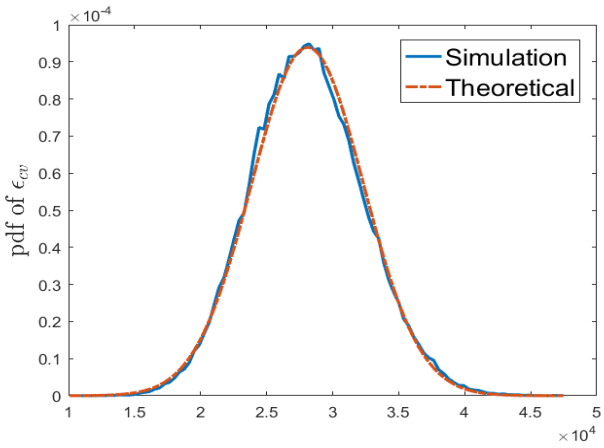

Fig. 3. The simulated (obtained over 10^5 noise realizations) and theoretically obtained pdf of $\epsilon_{cv,\ell_1,\lambda}$ (cf: Lemma 1) match closely.

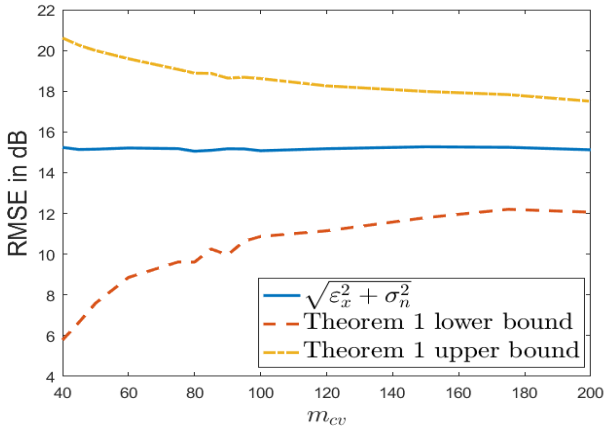

Fig. 4. Empirical demonstration of confidence intervals from Theorem 1.

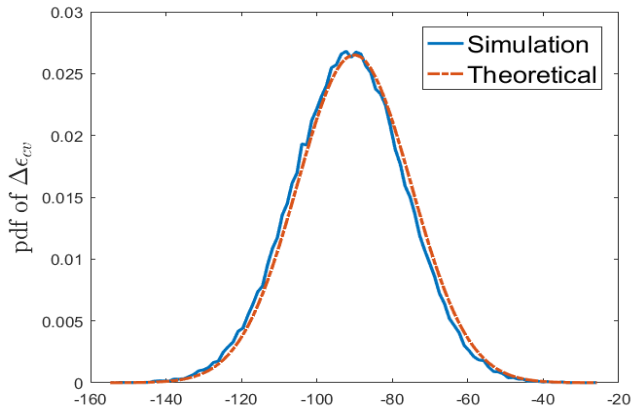

Fig. 5. The simulated (obtained over 10^5 noise realizations) and theoretically obtained pdf of $\Delta\epsilon_{cv} := \epsilon^p_{\text{cv}} - \epsilon^q_{cv}$ (cf: Lemma 2) match closely.

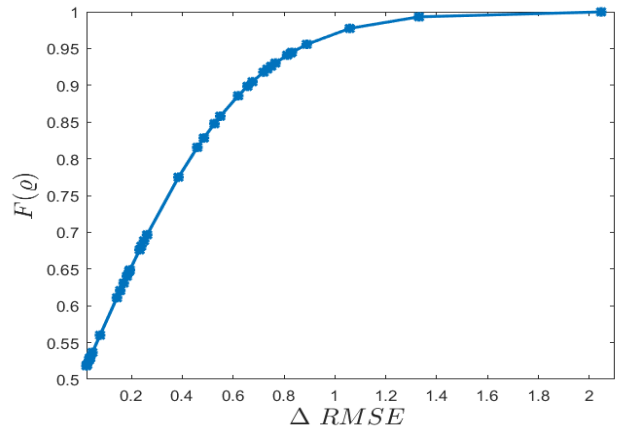

Fig. 6. Probability that $\epsilon^p_{cv} \geq \epsilon^q_{cv}$ as given by Theorem 2, plotted against $\Delta RMSE := \frac{|\varepsilon^p_x - \varepsilon^q_x|}{\|\boldsymbol{x}\|_2}$.

ables in each instance. This figure as well as Theorem 1 both predict that the bounds become tighter with increase in m_{cv}, which is quite intuitive.

Lemma 2: Let $\hat{\boldsymbol{x}}^p$ and $\hat{\boldsymbol{x}}^q$ be two recovered signals with their respective ℓ_1-CV errors $\epsilon^p_{cv}, \epsilon^q_{cv}$ and respective true recovery errors $\varepsilon_p, \varepsilon_q$. Define $\Delta\boldsymbol{x}^p := \boldsymbol{x} - \boldsymbol{x}^p, \Delta\boldsymbol{x}^q := \boldsymbol{x} - \boldsymbol{x}^q$. Assuming that $\mu_g \gg \sigma_g, \sigma_n, \varepsilon_p, \varepsilon_q$ and that m_{cv} is sufficiently large, we have $\Delta\epsilon_{cv} := \epsilon^p_{\text{cv}} - \epsilon^q_{cv} \sim \mathcal{N}\left(\mu, \sigma^2\right)$ where

$$\begin{aligned}
\mu =& (1-b)m_{cv}K_1\left(\sigma_p - \sigma_q\right)\\
\sigma^2 =& (1-b)m_{cv}\left(\sigma_p^2 + \sigma_q^2 - 2\rho_1\sigma_p\sigma_q\right) + m_{cv}\frac{b}{m}\left(\varepsilon_p^2 + \varepsilon_p^2\right)\\
&- \frac{2bm_{cv}}{m}\langle\Delta x^p, \Delta x^q\rangle - m_{cv}((1-b)K_1(\sigma_p - \sigma_q))^2\\
\rho_1 =& \frac{\sigma_p\sigma_q}{\pi}\left(\pi\rho_2 - 2\rho_2\tan^{-1}\left(\frac{\sqrt{1-\rho_2^2}}{\rho_2}\right) + 2\sqrt{1-\rho_2^2}\right)\\
\rho_2 =& \frac{\sigma_n^2 + \langle\Delta\mathbf{x}^p, \Delta\mathbf{x}^q\rangle}{m\sigma_p\sigma_q}\\
\sigma_p =& \sqrt{\frac{\varepsilon_p^2 + \sigma_n^2}{m}}, \qquad \sigma_q = \sqrt{\frac{\varepsilon_q^2 + \sigma_n^2}{m}}, \qquad K_1 = \sqrt{\frac{2}{\pi}}.
\end{aligned}$$

■

The proof of Lemma 2 can be found in the supp. material [10]. We demonstrate the result of Lemma 1 in Fig. 5 for $N = 1200, m = 440, m_{cv} = 40, b = 0.1, \sigma_n = 0.5, \sigma_g = 100, \mu_g = 700$. We conduct some experiments similar to the ones in Fig. 5 which demonstrate that the distribution of $\Delta\epsilon_{\text{cv}}$ obtained in Lemma 2 indeed matches very closely with empirical results corroborating our assumptions.

The following theorem shows that if the ℓ_1 CV error of one recovered signal is larger than that of another recovered signal, then with high probability, the ℓ_2 recovery errors for those signals follow the same order.

Theorem 2: Let $\hat{\boldsymbol{x}}^p$ and $\hat{\boldsymbol{x}}^q$ be two recovered signals, with (unobservable) recovery errors $\varepsilon^p_x, \varepsilon^q_x$ and corresponding cross-validation errors $\epsilon^p_{cv}, \epsilon^q_{cv}$. Assume $\mu_g >> \sigma_g, \sigma_n, \varepsilon_p, \varepsilon_q$, and m_{cv} is sufficiently large. If $\varepsilon^p_x \geq \varepsilon^q_x$, then it holds with probability $F(\varrho)$ that $\epsilon^p_{cv} \geq \epsilon^q_{cv}$, where $\varrho = \frac{\mu}{\sigma}$ where, μ, σ are as defined in Lemma 2 and F is the standard Gaussian CDF. ■

From the expressions of μ, σ, ϱ, one can observe that the confidence with which $\varepsilon^p_{\text{x}} \geq \varepsilon^q_{\text{x}}$ holds increases monotonically with m_{cv} and tends to 1 as m_{cv} increases. This is because, with all other parameters fixed, $\frac{\mu}{\sigma}$ is proportional to $\sqrt{m_{cv}}$. In Fig 6 we demonstrate this idea for $N = 1200, m = 420, m_{cv} = 20, \|\boldsymbol{x}\|_0 = 50, b = 0.05, \sigma_n = 0.5, \mu_g = 1000, \sigma_g = 20$. It is evident from Fig.6 that, as the difference in the cross validation error of two recovered signals increases, the likelihood that their recovery error follows the same order increases. This observation is very crucial as it strongly supports the idea that ℓ_1-CV error is a very good metric for choosing the optimal regularization parameter in our experiments.

4. CONCLUSION

We provide a detailed theoretical and empirical analysis of the ℓ_1-CV error when used for compressive reconstruction in presence of mixed impulse and Gaussian noise. Under some assumptions, we prove that the ℓ_1 CV error follows a Gaussian distribution and further provide upper and lower bounds on the recovery error of a signal estimate in terms of the ℓ_1-CV error and perform simulations for generic parameters to validate these results. We justify the use of ℓ_1-CV error for selecting the optimal parameter in signal reconstruction algorithms like ℓ_1-LASSO, by proving that with high probability, the ordering of the actual recovery error of any two signal estimates is the same as the ordering of their ℓ_1-CV errors.

5. REFERENCES

[1] E. Candes, "The restricted isometry property and its implications for compressive sensing," *Comptes Rendus Mathematiques*, 2008.

[2] T. Hastie, R. Tibshirani, and M. Wainwright, *Statistical Learning with Sparsity: The LASSO and Generalizations*, CRC Press, 2015.

[3] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[4] P. Boufounos, M. F. Duarte, and R. G. Baraniuk, "Sparse signal reconstruction from noisy compressive measurements using cross validation," in *IEEE SSP*, 2007, pp. 299–303.

[5] R. Ward, "Compressed sensing with cross validation," *IEEE Transactions on Information Theory*, vol. 55, no. 12, 2009.

[6] W. Johnson and J. Lindenstrauss, "Extensions of lipschitz maps into a hilbert space," *Contemp. Math*, vol. 26, 1984.

[7] J. Zhang, L. Chen, P. T. Boufounos, and Y. Gu, "On the theoretical analysis of cross validation in compressive sensing," in *ICASSP*, 2014, pp. 3370–3374.

[8] G. Tzagkarakis, J. P. Nolan, and P. Tsakalides, "Compressive sensing using symmetric alpha-stable distributions for robust sparse signal reconstruction," *IEEE Trans. Signal Process.*, vol. 67, no. 3, pp. 808–820, 2019.

[9] R. Carrillo et al, "Robust compressive sensing of sparse signals: a review," *EURASIP J. Adv. Signal Process.*, vol. 108, 2016.

[10] "Supplemental material," `https://www.cse.iitb.ac.in/~ajitvr/ICASSP2021_supp/CV_supp.pdf`.

[11] "Folded normal distribution," `https://en.wikipedia.org/wiki/Folded_normal_distribution`.

[12] F. C. Leone, L. S. Nelson, and R. B. Nottingham, "The folded normal distribution," *Technometrics*, vol. 3, no. 4, pp. 543–550, 1961.