SPARSE CHANNEL ESTIMATION WITH REGULARIZATION METHODS IN MASSIVE MIMO SYSTEMS

Ture Peken, Ravi Tandon, and Tamal Bose
University of Arizona
Tucson, AZ, 85719
[turepeken,tandonr,tbose]@email.arizona.edu

ABSTRACT

Massive multiple-input multiple-output (MIMO) technology has recently gained a lot of attention as a candidate technology for the next generation wireless systems. With a higher number of antennas, pilot-based channel estimation faces a limitation in the number of orthogonal pilots to be used among users in all cells. Sparse channel estimation by using regularization methods can reduce the pilots compared to pilot-based channel estimation. In this paper, we study two regularization methods: least absolute shrinkage and selection operator (lasso) and elastic net. We investigate the performance of least squares (LS), lasso, and elastic net when the sparsity of the channel changes over time. We study the optimum tuning parameters for lasso and elastic net based channel estimators to achieve the best performance with the different number of pilots and values of signal-to-noise ratio (SNR). Finally, we present the asymptotic analysis of LS, lasso, and elastic net based channel estimators.

Keywords - Sparse channel estimation, massive MIMO, lasso, elastic net

INTRODUCTION

Multiple antenna technology is a basic feature of all advanced wireless communication systems such as IEEE 802.16M and 3GPP LTE/LTE-Advanced [1]. Next generation wireless systems leverage a large excess of antennas at the base station (BS), which is referred as massive MIMO, to provide better spectral and energy efficiency, higher data rates and capacity [2]. In massive MIMO systems, BSs are equipped with a number of antennas much larger than the UTs in the same frequency-time resource, which improves the spectral efficiency compared to conventional MIMO. Energy efficiency can be obtained by adding the signals sent from the antennas of the BS constructively at the desired UTs and destructively at the remaining UTs.Therefore, massive MIMO is considered to be one of the key technologies for fifth generation (5G) wireless networks.

Evaluation of channel state information (CSI) with minimum error is very important to recover received symbols correctly in a wireless communication system. Currently, pilot-based channel estimation methods are used in most advanced wireless systems. CSI is estimated by transmitting a set of pilot sequences from each UT to the BS in pilot-based methods. In massive MIMO systems, one of the main challenges is the interference among pilots in different cells. This phenomenon, which is called pilot contamination, occurs due the reuse of orthogonal pilot sequences in different cells. When multiple UTs share same pilot sequences, interference between these UTs is inevitable [2]. Therefore, pilot-based channel estimation methods limit the gain that can be achieved with massive MIMO.

The present paper investigates the usage of shorter pilot sequences to decrease the pilot contamination effect and the overhead due to transmission of pilots. Since the channel response between the BS and the UT becomes sparse as the number of the antennas at the BS increases, sparse channel estimation techniques can be leveraged to shorten the pilot sequences [3]. In [4, 5, 6], sparsity of massive MIMO channel is exploited by using different techniques to reduce the number of symbols in pilot sequences. In this paper, we study sparse channel estimation by using two regularization methods: lasso and elastic net. Lasso is a L_1-penalized LS method which provides a good estimation for sparse channels [7]. Most of the sparse channel estimation methods in the literature are based on L_1 penalization [3, 8, 9]. Elastic net, which uses both L_1 and L_2 penalization, is another popular regularization method that does the variable selection and shrinkage simultaneously [9]. Elastic net improves the prediction accuracy compared to lasso when the pairwise correlations among the group of variables are high.

In [12], we propose an elastic net based channel estimation for massive MIMO to overcome pilot contamination effect by decreasing the length of pilot sequences. In this work, we compare root mean square error (RMSE) and computational complexity of elastic net based channel estimation with LS and lasso by using MATLAB simulations. We observe the performance of these methods with and without pilot contamination while number of pilots and values of SNR change. This paper's main contributions are:

1. We study the effect of channel sparsity on the performance of LS, lasso, and elastic net based channel estimation methods when the different number of pilots and values of SNR are used.

2. We investigate the selection of optimum tuning parameters for lasso and elastic net based channel estimators.

3. We analyze the asymptotic performance of LS, lasso, and elastic net based channel estimators when the length of pilot sequences is in large regime.

In this paper, performance of LS, lasso, and elastic net with different levels of channel sparsity are compared in terms of RMSE by using MATLAB simulations. It is shown that elastic net gives the least RMSE when the sparsity of channel is high. Moreover, lasso and elastic net achieve similar RMSE while channel sparsity decreases. Optimum tuning parameters are evaluated for lasso and elastic net by using cross-validation. According to our asymptotic analysis, it is shown that the variance of lasso and elastic net channel estimators increases when SNR decreases and K approaches to ∞. We also show with our asymptotic analysis that the bias of elastic net and lasso channel estimators becomes significantly large when magnitude of channel coefficients are large.

SYSTEM MODEL

We consider a massive MIMO system in which BS has M antennas and each UT has a single antenna. It is assumed that there are N UTs where $M \gg N$. In each frame, K number of symbols are transmitted as a pilot sequence. The L received symbols at the M antennas of the BS are given as:

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{W}, \tag{1}$$

where $\mathbf{Y}\in\mathbb{C}^{M\times L}$ is a complex valued received symbols matrix, $\mathbf{X}\in\mathbb{C}^{N\times L}$ is a complex valued matrix of L transmitted symbols. $\mathbf{H}\in\mathbb{C}^{M\times N}$ is the CSI between N UTs and M antennas of the BS, and $\mathbf{W}\in\mathbb{C}^{L\times M}$ is the additive white Gaussian noise (AWGN) matrix, i.e. $\mathcal{N}(\mathbf{0}, \sigma^2\boldsymbol{I})$.

Radio channel in massive MIMO consists of different multipath components (MPCs), which are caused due to the interaction between the radio waves and the scatterers in the environment. Cluster-based channels are proposed to reduce the number of channel modeling parameters by grouping MPCs with the same delay and directions into clusters. Therefore, we extend a cluster-based channel model (COST 2100) in this paper. COST 2100 captures important massive MIMO channel characteristics [10].

CHANNEL ESTIMATION

In conventional MIMO systems, channel estimation is done by transmitting known pilot sequences from the UTs to the BS. LS channel estimation method is one of the most common pilot-based channel estimation methods due to its low computational complexity. If the same pilot sequences are used in R cells, i.e., $\mathbf{X}_{p_1} = ... = \mathbf{X}_{p_R}$. The pilot sequence of the ith cell is denoted as $\mathbf{X}_{p_i}$. Then, the LS estimator for channel matrix between the BS and the ith cell is given by,

$$\hat{\mathbf{H}}_i = (\mathbf{X}_{p_i}^H\mathbf{X}_{p_i})^{-1}\mathbf{X}_{p_i}\mathbf{Y} = \mathbf{H}_i + \sum_{j\neq i, j=1}^{R}\mathbf{H}_j + (\mathbf{X}_{p_i}^H\mathbf{X}_{p_i})^{-1}\mathbf{X}_{p_i}\mathbf{W}. \tag{2}$$

It is seen in (2) that the interfering channels H_j, $j = 1, ..., R$, $j \neq i$ will contaminate the desired channel estimate H_i. This effect is called pilot contamination [11].

SPARSE CHANNEL ESTIMATION

A channel is considered as sparse if number of multipath components is smaller than length of the channel. Sparsity of channel can be measured by different metrics. L_p norm-like measures are widely used as sparsity measures,

$$||\mathbf{v}||_p = \left(\sum_i {v_i}^p\right)^{\frac{1}{p}}. \tag{3}$$

In this paper, sparsity of the channel is measured by using L_1 norm, which is a form of regularization. Suppose that an outcome vector $\mathbf{y} \in \mathbb{R}^n$ and a predictor matrix $\mathbf{X} \in \mathbb{R}^{n \times p}$, whose columns $\mathbf{x}_1, ..., \mathbf{x}_p \in \mathbb{R}$ denote predictor variables, are observed. Lasso regression solves LS problem by constraining the coefficients by their L_1 norm:

$$\min_{\boldsymbol{\beta} \in \mathbb{R}^p} \sum_{i=1}^{n} \left(\boldsymbol{y}_i - \sum_{j=1}^{p} \beta_j \mathbf{X}_j\right)^2 \qquad \text{subject to} \qquad ||\boldsymbol{\beta}||_1 \leq \lambda, \tag{4}$$

where λ is a fixed non-negative tuning parameter. By using L_1 norm, we restrict our estimate to lie in a ball around 0. This brings sparsity to the estimate.

Channel estimation based on lasso applies a lasso regression on each column vector of the complex channel matrix. In this case, the jth column vector of the complex channel matrix $\mathbf{H}$ between N UTs and jth antenna of the BS is given as,

$$\hat{\mathbf{h}}_j = \arg\min_{\mathbf{h}_j} ||\mathbf{y}_j - \mathbf{X}_p \mathbf{h}_j||_2^2 + \lambda ||\mathbf{h}_j||_1, \tag{5}$$

where $\mathbf{X}_p$ are the pilot sequences such that $\mathbf{X}_p = (\mathbf{x}_1|\mathbf{x}_2|...|\mathbf{x}_N)$ and $\mathbf{x}_i = (x_{i_1}, x_{i_2}, ..., x_{i_K})$, $i = 1, ..., N$ are the predictors. $\mathbf{y}_j$ is the vector of K received pilots at the jth antenna of the BS.

Even though lasso is more powerful for estimating sparse channels compared to LS, it has some limitations. For $K > N$, if correlations are high between predictors, the prediction performance of lasso starts to degrade. Another regularization method, called elastic net, has been proposed to overcome the limitations of lasso. Elastic net regularization solves LS problem by constraining the coefficients by both their L_1 and L_2 norms:

$$\min_{\boldsymbol{\beta} \in \mathbb{R}^p} \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{p} \beta_j \mathbf{X}_j\right)^2 \qquad \text{subject to} \qquad ||\boldsymbol{\beta}||_2 \leq \lambda_2 \qquad \text{and} \qquad ||\boldsymbol{\beta}||_1 \leq \lambda_1 \tag{6}$$

where λ_1 and λ_2 are fixed non-negative upper bounds on L_1 and L_2 norms, respectively.

Channel estimation based on elastic net has been proposed in [12]. The jth column vector of the complex channel matrix $\mathbf{H}$ between N UTs and jth antenna of the BS is given as:

$$\hat{\mathbf{h}}_j = \left(1 + \frac{\lambda_2}{n}\right) \underset{\mathbf{h}_j}{\arg\min} \, ||\mathbf{y}_j - \mathbf{X}_p \mathbf{h}_j||_2^2 + \lambda_2 ||\mathbf{h}_j||_2^2 + \lambda_1 ||\mathbf{h}_j||_1. \tag{7}$$

Least angle regression (LARS) algorithm, which is proposed for elastic net regression in [13], is extended for lasso based channel estimation in this paper. At the beginning of this algorithm, all coefficients are initialized to zero. Then, the pilot sequence $\boldsymbol{x}_{j_1}$, which is most correlated with received symbols vector $\boldsymbol{y}_j$, is evaluated. Next, the largest possible step is taken in the direction of $\boldsymbol{x}_{j_1}$ until another pilot sequence $\boldsymbol{x}_{j_2}$ has the same correlation with the current residual. The algorithm continues along the least angle direction between $\boldsymbol{x}_{j_1}$, $\boldsymbol{x}_{j_2}$, and $\boldsymbol{x}_{j_3}$ until another pilot sequence earns its way to the active set which is the most correlated set, so on. LARS-EN algorithm is proposed for elastic net regression in [9]. The elastic net problem is equivalent to the lasso problem for a given λ_2. This algorithm is extended to elastic net based channel estimation in this work.

Calculating the right values of the tuning parameters are key to the performance of the lasso and elastic net algorithms. One of the primary methods for estimating a tuning parameter λ is K-fold cross-validation. In lasso, we need to to cross-validate on one-dimensional surface since there is one tuning parameter. On the other hand, we need to cross-validate on two-dimensional surface for elastic net. In this paper, we apply fivefold cross-validation. In elastic net, a grid of values are first selected for the first tuning parameter. Then, for each value of this tuning parameter, LARS-EN algorithm generates the solution path for the elastic net and applies cross-validation to select the second tuning parameter. Finally, the first tuning parameter is selected as the one with the smallest cross-validation error.

ASYMPTOTIC ANALYSIS FOR SPARSE CHANNEL ESTIMATION

In this section, we will use the following assumptions given in (8) and (9).

$$\lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} \boldsymbol{x}_k {\boldsymbol{x}_k}^T = \boldsymbol{C}, \tag{8}$$

and

$$\lim_{K\to\infty} \frac{1}{K} \max_{1\leq k\leq K} {\boldsymbol{x}_k}^T \boldsymbol{x}_k = \mathbf{0}. \tag{9}$$

Here, $\boldsymbol{C}$ is a nonnegative definite matrix, K is the length of pilot sequence in symbols, and $\boldsymbol{x}_k$ is the kth row vector of $\boldsymbol{X}_p$ transmitted pilot symbols matrix. Then, LS channel estimator is consistent and that [14]:

$$\sqrt{K}\left(\hat{\mathbf{h}}_j - \mathbf{h}_j\right) \to \mathcal{CN}(\mathbf{0}, \sigma^2 \boldsymbol{C}^{-1}) \approx \mathcal{CN}\left(\mathbf{0}, \frac{1}{SNR}\right) \tag{10}$$

For asymptotic analysis of lasso, let us define a tuning parameter $\lambda_K = o(K)$ when pilot sequences with length of K symbols are used. By using Theorem 2 in [14], if $\frac{\lambda_K}{\sqrt{K}} \rightarrow \lambda_0 \geq 0$ and $\boldsymbol{C}$ is nonsingular, lasso channel estimator satisfies the following:

$$\sqrt{K}\left(\hat{\mathbf{h}}_j - \mathbf{h}_j\right) \xrightarrow{d} \underset{\boldsymbol{t}\in\mathbb{C}^{N\times 1}}{\arg\min}(D(\boldsymbol{t})), \tag{11}$$

where

$$D(\boldsymbol{t}) = -2\boldsymbol{t}^T\boldsymbol{G} + \boldsymbol{t}^T\boldsymbol{C}\boldsymbol{t} + \lambda_0 \sum_{n=1}^{N} \left[t_n sgn(h_{jn}) I(h_{jn} \neq 0) + |t_n| I(h_{jn} = 0)\right]. \tag{12}$$

Here, arg min denotes the value of the argument $\boldsymbol{t}$ that minimizes the objective function $D(\cdot)$, $\xrightarrow{d}$ represents convergence in distribution, t_n is the nth element of the vector $\boldsymbol{t}$, N is the number of UTs, h_{jn} is the channel coefficient between jth antenna of the BS and nth UT in the massive MIMO system, and $\boldsymbol{G} \in \mathbb{C}^{N\times N}$ has a $\mathcal{CN}(\mathbf{0}, \sigma^2\boldsymbol{C})$ distribution. Note that when $\lambda_0 = 0$ and $K \rightarrow \infty$, $\arg\min_{\boldsymbol{t}\in\mathbb{C}^{N\times 1}}(D(\boldsymbol{t})) = \boldsymbol{C}^{-1}\boldsymbol{G} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\boldsymbol{C}^{-1}) \approx \mathcal{CN}(\mathbf{0}, \frac{1}{SNR})$.

Let us define tuning parameters of elastic net as λ_1^K and λ_2^K when pilot sequences with length of K symbols are used. It is considered that $\lambda_1^K = o(K)$ and $\lambda_2^K = o(K)$ for asymptotic analysis of elastic net. We also assume that $\frac{\lambda_1^K}{\sqrt{K}} \rightarrow \lambda_1^0 \geq 0$, $\frac{\lambda_2^K}{\sqrt{K}} \rightarrow \lambda_2^0 \geq 0$, and $\boldsymbol{C}$ is nonsingular. From Theorem 2 in [14], elastic net channel estimator is consistent such that:

$$\sqrt{K}\left(\hat{\mathbf{h}}_j - \mathbf{h}_j\right) \xrightarrow{d} \underset{\boldsymbol{u}\in\mathbb{C}^{N\times 1}}{\arg\min}(V(\boldsymbol{u})), \tag{13}$$

where

$$\boldsymbol{V}(\boldsymbol{u}) = -2\boldsymbol{u}^T\boldsymbol{W} + \boldsymbol{u}^T\boldsymbol{C}\boldsymbol{u} + \lambda_1^0 \sum_{n=1}^{N} \left[u_n sgn(h_{jn}) I(h_{jn} \neq 0) + |u_n| I(h_{jn} = 0)\right] + \lambda_2^0 \sum_{n=1}^{N} u_n sgn(h_{jn})|h_{jn}|. \tag{14}$$

Here, arg min denotes the value of the argument $\boldsymbol{u}$ that minimizes the objective function $V(\cdot)$, $\xrightarrow{d}$ represents convergence in distribution, u_n is the nth element of the vector $\boldsymbol{u}$, $\boldsymbol{W} \in \mathbb{C}^{N\times N}$ has a $\mathcal{CN}(\mathbf{0}, \sigma^2\boldsymbol{C})$ distribution and h_{jn} denotes the channel coefficient between jth antenna of the BS and nth UT in the system. Note that $\sqrt{K}\left(\hat{\mathbf{h}}_j - \mathbf{h}_j\right) \xrightarrow{d} \boldsymbol{C}^{-1}(2\boldsymbol{W} - \lambda_2^0\mathbf{h}_j) \sim \mathcal{CN}(-\lambda_2^0\boldsymbol{C}^{-1}\mathbf{h}_j, \sigma^2\boldsymbol{C}^{-1}) \approx \mathcal{CN}(-\lambda_2^0\boldsymbol{C}^{-1}\mathbf{h}_j, \frac{1}{SNR})$ for elastic net in the large regime of K.

Asymptotic results indicate that we need $\lambda_K = O(\sqrt{K})$ while $K \rightarrow \infty$ for a consistent lasso channel estimator. In order to achieve a consistent elastic net channel estimator, we require $\lambda_1^K = O(\sqrt{K})$ and $\lambda_2^K = O(\sqrt{K})$ according to asymptotic analysis. Moreover, variance of lasso and elastic net channel estimator increases while SNR decreases in the large regime of K. These results also illustrate that the amount of shrinkage with lasso and elastic net towards 0 increases with the magnitude of channel coefficient vector $\mathbf{h}_j$. Therefore, the bias of lasso and elastic net channel estimators can be significantly large for channels with low level of sparsity.

SIMULATION RESULTS

In this section, a multi-cellular massive MIMO system is simulated in MATLAB. In this system, each BS has 100 antennas and different number of single-antenna UTs in its coverage. The total number of UTs in the system equals to 90. The channel is modeled as COST 2100 with Rayleigh fading. We consider a carrier frequency of 2.4 GHz for this massive MIMO system.

First, we observe the performance of LS, lasso, and elastic net while the sparsity of channel changes over time. Figure 1 and 2 show RMSE between real and estimated complex channel matrices of LS, lasso and elastic net methods for different values of λ_1 when number of pilots are 20 and 12, respectively. In both figures, SNR is 10 dB. It can be seen that, performance of all the methods get worse as the sparsity of the channel increases. When number of pilots is significantly small, lasso, and elastic net outperform pilot-based channel estimation methods such as LS. On the other hand, LS achieves smaller channel estimation error than the sparse channel estimation methods while number of pilots increases.

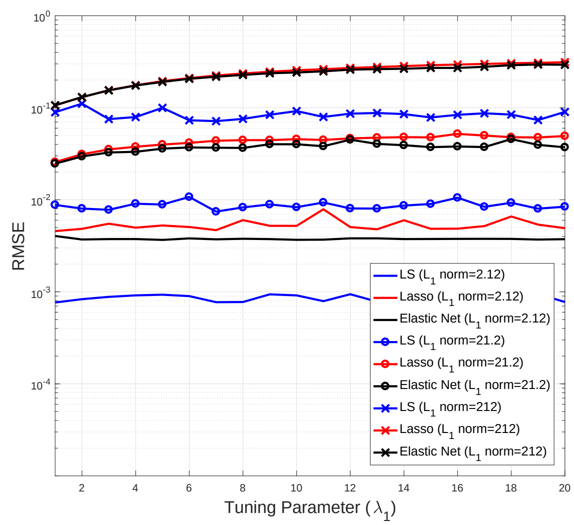

Figure 1: RMSE between real and estimated complex channel matrices of LS, lasso, and elastic net methods for different values of λ_1. Number of pilots and data symbols are 20 and 140, respectively.

We also compare the performance change with sparsity for the different number of pilots. For these results, tuning parameters for elastic net and lasso were chosen for the different number of pilots. For lasso, fivefold cross validation is applied for $\lambda_1 = [0, 0.01, 0.1, 1, 10, 100]$ when number of pilots increase from 12 to 30. In elastic net, a grid values for λ_1 is selected as $\lambda_1 = [0, 0.01, 0.1, 1, 10, 100]$. Then, for each value of λ_1, fivefold cross-validation is applied to select the λ_2. λ_1, which gives the smallest cross-validation error, is chosen. Based on the simulations, $\lambda_1 = \lambda_2 = 0.1$ are selected. Figure 3 show RMSE between real and estimated complex channel matrices of LS, lasso and elastic net for channels with different levels of sparsity while number of pilots increases from 12 to 30. Here, SNR is 20 dB and total number of symbols equals to 160. It can be seen that elastic net outperforms lasso and LS when

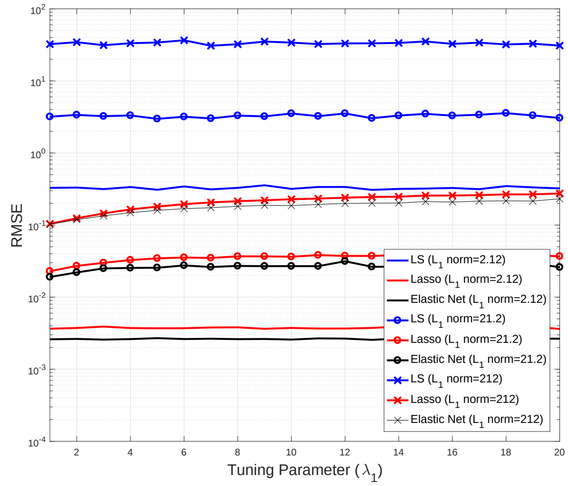

Figure 2: RMSE between real and estimated complex channel matrices of LS, lasso, and elastic net methods for different values of λ_1. Number of pilots and data symbols are 12 and 148, respectively.

number of pilots is less than 20 for channels with different levels of sparsity. Lasso achieves slightly worse RMSE than elastic net but outperforms LS when number of pilots is small.

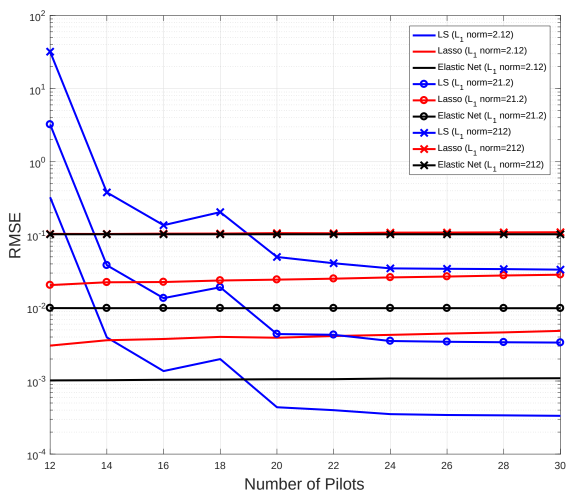

Figure 3: RMSE between real and estimated complex channel matrices of LS, lasso and elastic net methods for different number of pilots and sparsity of channels.

Finally, we observe how RMSE changes with increasing values of SNR when channels with different levels of sparsity are used. Figure 4 shows RMSE of LS, lasso, and elastic net based channel estimation methods while SNR increases from -20 to 20 dB. Here, results are obtained when number of pilot and data symbols are 12 and 148, respectively.

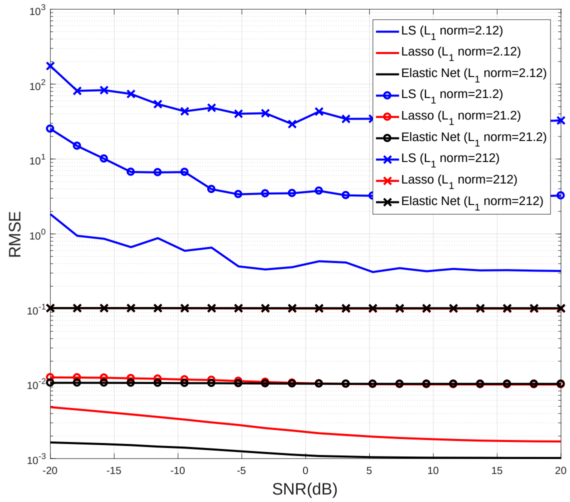

Figure 4: RMSE between real and estimated complex channel matrices of LS, lasso and elastic net methods for different values of SNR and sparsity of channels.

CONCLUSIONS

Massive MIMO is expected to be one of the key technologies in the next generation wireless communications systems. Pilot-based channel estimation methods bring significant overhead and cause pilot contamination due to the upper bound on the number of orthogonal pilots to be used for the users in different cells. In this paper, sparsity of massive MIMO channel is exploited by using regularization methods: lasso and elastic net to estimate the channel with less number of pilots compared to pilot-based channel estimation methods. Asymptotic analysis for LS, lasso, and elastic net channel estimators is also presented. With asymptotic analysis, we show that variance of lasso and elastic net channel estimator increases with decreasing values of SNR in the large regime of K. Furthermore, the bias of lasso and elastic net channel estimators can be unacceptably large when channel has a low sparsity. Performance of these methods are compared with a popular pilot-based channel estimation method: LS by using MATLAB simulations. By using cross-validation, optimum tuning parameters are obtained for lasso and elastic net. For channels with different levels of sparsity, performance of LS, lasso, and elastic net are compared while number of pilots and values of SNR change. It is seen that elastic net outperforms both lasso and LS when sparsity of channel is high. However, lasso and elastic net start to achieve similar performance while sparsity of channel decreases.

ACKNOWLEDGEMENTS

This project was partially supported by the Broadband Wireless Access and Applications Center (BWAC); NSF Award No. 1265960.

REFERENCES

[1] Q. Li, G. Li, W. Lee, M. il Lee, D. Mazzarese, B. Clerckx, and Z. Li, "MIMO Techniques in WiMAX and LTE: A Feature Overview," *Communications Magazine, IEEE*, vol. 48, pp. 86–92, May 2010.

[2] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for Next Generation Wireless Systems," *Communications Magazine, IEEE*, vol. 52, pp. 186–195, February 2014.

[3] G. Destino, M. Juntti, and S. Nagaraj, "Leveraging sparsity into massive MIMO channel estimation with the adaptive-LASSO," in *2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 166–170, Dec 2015.

[4] J. W. Choi, B. Shim, and S. H. Chang, "Downlink Pilot Reduction for Massive MIMO Systems via Compressed Sensing," *IEEE Communications Letters*, vol. 19, pp. 1889–1892, Nov 2015.

[5] M. Masood, L. H. Afify, and T. Y. Al-Naffouri, "Efficient Coordinated Recovery of Sparse Channels in Massive MIMO," *IEEE Transactions on Signal Processing*, vol. 63, pp. 104–118, Jan 2015.

[6] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed Channel Sensing: A New Approach to Estimating Sparse Multipath Channels," *Proceedings of the IEEE*, vol. 98, pp. 1058–1076, June 2010.

[7] R. Tibshirani, "Regression Shrinkage and Selection Via the Lasso," *Journal of the Royal Statistical Society, Series B*, vol. 58, pp. 267–288, 1994.

[8] J. W. Choi and B. Shim, "Compressive sensing based pilot reduction technique for massive MIMO systems," in *2015 Information Theory and Applications Workshop (ITA)*, pp. 115–118, Feb 2015.

[9] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 67, no. 2, pp. 301–320, 2005.

[10] L. Liu, C. Oestges, J. Poutanen, K. Haneda, P. Vainikainen, F. Quitin, F. Tufvesson, and P. Doncker, "The COST 2100 MIMO Channel Model," *Wireless Communications, IEEE*, vol. 19, pp. 92–99, December 2012.

[11] T. Marzetta, "Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas," *Wireless Communications, IEEE Transactions on*, vol. 9, pp. 3590–3600, November 2010.

[12] T. Peken, R. Tandon, and T. Bose, "Elastic net for channel estimation in massive MIMO," *Proceedings of the International Telemetering Conference*, 1 2017.

[13] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Ann. Statist.*, vol. 32, pp. 407–499, 04 2004.

[14] K. Knight and W. Fu, "Asymptotics for lasso-type estimators," *Ann. Statist.*, vol. 28, pp. 1356–1378, 10 2000.