2014 IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)

# ROBUST REGULARIZED LEAST SQUARES ESTIMATION IN THE PRESENCE OF BOUNDED DATA UNCERTAINTIES

*N. Denizcan Vanli, Suleyman S. Kozat*

Bilkent University
Ankara, Turkey

*Mehmet A. Donmez*

University of Illinois at Urbana-Champaign
IL, USA

**ABSTRACT**

We study the problem of estimating an unknown deterministic signal that is observed through an unknown deterministic observation matrix and additive noise under the regularized residual error criterion. In this framework, we introduce a robust approach to this problem and consider the performance of an estimator relative to the performance of the least squares (LS) estimator tuned to the underlying unknown observation matrix and noise. This relative performance measure in fact turns out to be the regret of the estimator for not knowing the true parameters. Refraining from any statistical and structural assumptions both on the observation matrix and noise, we then minimize this regret with a robust LS estimation method, where we also demonstrate that this method can be cast as a semi-definite programming (SDP) problem. Numerical examples are also presented to illustrate the theoretical results.

***Index Terms*—** Data estimation, least squares, regularized, robust, minimax regret.

## 1. INTRODUCTION

In this paper, we consider a linear estimation problem from a competitive framework under the regularized residual error criterion. Here, an unknown deterministic signal is observed through an unknown deterministic observation matrix, where the output of the system is corrupted by additive noise. Although the observation matrix and the output vector are not exactly known, estimates for both of them as well as the uncertainty bounds are given [1]. Under such a scenario, using a direct regularized LS estimator tuned to the estimations may significantly deteriorate the performance especially when the perturbations on the observation matrix and the output vector are relatively high [1]. In order to cope with such performance degradations, a common method is to use the robust regularized LS method, which minimizes the residual error with respect to the worst-case perturbations [2]. Although such methods minimize the residual error for the worst-case perturbations, they usually provide unsatisfactory results in the average. Therefore, in order to counterbalance the conservative nature of the robust regularized LS methods, we introduce a competitive approach to this problem and define the performance of an estimator relative to the performance of the least squares (LS) estimator tuned to the underlying unknown observation matrix and noise. This relative performance difference is called as "regret" [3–5] of not knowing the true parameters. In this sense, we introduce a novel robust regularized LS method minimizing this worst-case regret and provide a performance trade-off between the robust regularized LS methods and the optimal regularized LS method tuned to the estimates.

There exists a wide range of applications in the signal processing literature that deals with LS problems [1, 4, 6]. However, the performance of the LS estimators may considerably degrade due to the possible errors in the observation matrix and the output vector in various applications. An appealing approach to find robust solutions to such estimation problems is the robust LS method [1], in which the uncertainties in the observation matrix and the output vector are incorporated into optimization framework via a minimax residual formulation. However, robust LS method is a pessimistic approach since it considers the worst possible perturbation under the uncertainty bounds. To alleviate the pessimistic nature of the worst-case optimization methods, minimax regret approaches have been presented in the signal processing literature [3–5].

In this paper, we introduce a competitive approach to the robust regularized LS problem and present a robust estimator whose performance is as close as possible to the one of the optimal estimator for all possible perturbations within the uncertainty bounds. We emphasize that our cost formulation differs from [3–5], where we consider a slightly more general loss function which often arises in many signal processing problems [2]. We also note that the competitive problem formulation of this paper is significantly different than [4, 5], where the regret term is directly used in the cost function. Although a similar regret notion is used in [3], not only the cost functions but also the problem formulation regarding the constraints on the uncertainties substantially differ in this paper.

The organization of the paper is as follows. An overview to the problem is provided in Section 2. In Section 3, we then introduce the robust regularized LS estimation method based on our regret formulation, and then provide the explicit SDP formulation for the problem. The numerical examples are demonstrated in Section 4 and the paper concludes with certain remarks in Section 5.

## 2. PROBLEM FORMULATION

### 2.1. Notation

In this paper, all vectors are column vectors and represented by boldface lowercase letters. Matrices are represented by boldface uppercase letters. For a matrix $\mathbf{H}$, $\mathbf{H}^H$ is the conjugate transpose, $||\mathbf{H}||$ is the spectral norm, $\mathbf{H}^+$ is the pseudo-inverse, $\mathbf{H} > 0$ represents a positive definite matrix and $\mathbf{H} \geq 0$ represents a positive semi-definite matrix. For a square matrix $\mathbf{H}$, $\mathrm{Tr}(\mathbf{H})$ is the trace. Naturally, for a vector $\mathbf{x}$, $||\mathbf{x}|| = \sqrt{\mathbf{x}^H\mathbf{x}}$ is the $L_2$-norm. Here, $\mathbf{0}$ denotes a vector or matrix with all zero elements and the dimensions can be understood from the context. Similarly, $\mathbf{I}$ represents the appropriate sized identity matrix. The operator $\mathrm{vec}(\cdot)$ is the vectorization operator, i.e., it stacks the columns of a matrix of dimension $m \times n$ into a $mn \times 1$ column vector. Finally, the operator $\otimes$ is the Kronecker product [7].

### 2.2. Problem Description

We investigate the problem of estimating an unknown deterministic vector $\mathbf{x} \in \mathbb{C}^n$, which is observed through an unknown deterministic observation matrix. However, instead of the actual observation matrix and the output vector, their estimates $\mathbf{H} \in \mathbb{C}^{m\times n}$ and $\mathbf{y} \in \mathbb{C}^m$ and uncertainty bounds on these estimates are provided, where $m \geq n$. Our aim is to find a solution to solve the regularized LS problem

$$\mathbf{y} \approx \mathbf{H}\mathbf{x},$$

such that

$$\mathbf{y} + \Delta\mathbf{y} = (\mathbf{H} + \Delta\mathbf{H})\mathbf{x},$$

for deterministic perturbations $\Delta\mathbf{H} \in \mathbb{C}^{m\times n}$, $\Delta\mathbf{y} \in \mathbb{C}^m$, where $\mathbf{H}+\Delta\mathbf{H}$ is a full rank matrix. Although these perturbations are unknown, a bound on each perturbation is provided, i.e.,

$$||\Delta\mathbf{H}|| \leq \delta_H \text{ and } ||\Delta\mathbf{y}|| \leq \delta_Y,$$

where $\delta_H, \delta_Y \geq 0$. In this sense, we refrain from any statistical and structural assumptions on the observation matrix and the output vector, yet consider that the estimates $\mathbf{H}$ and $\mathbf{y}$ are at least accurate to "some degree" but their actual values under these uncertainties are completely unknown to the observer. In this framework, we define the loss function as the regularized residual error, i.e.,

$$l(\mathbf{x}; \Delta\mathbf{H}, \Delta\mathbf{y}) \triangleq ||(\mathbf{y} + \Delta\mathbf{y}) - (\mathbf{H} + \Delta\mathbf{H})\mathbf{x}||^2 + \mu\,||\mathbf{x}||^2, \tag{1}$$

where $\mu > 0$ is the regularization parameter.

Even in the presence of these uncertainties, the symbol vector $\mathbf{x}$ can be naively estimated by simply substituting the estimates $\mathbf{H}$ and $\mathbf{y}$ into the optimal regularized LS estimator [2] and obtain the following solution

$$\hat{\mathbf{x}} = \left(\mathbf{H}^H\mathbf{H} + \mu\mathbf{I}\right)^{-1}\mathbf{H}^H\mathbf{y}.$$

However, this approach yields unsatisfactory results, when the errors in the estimates of the channel matrix and the output vector are relatively high [1, 4–6]. A common approach to find a robust solution is to employ a worst-case regularized residual minimization [1]

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbb{C}^n} \max_{||\Delta\mathbf{H}||\leq\delta_H, ||\Delta\mathbf{y}||\leq\delta_Y} l(\mathbf{x}; \Delta\mathbf{H}, \Delta\mathbf{y}),$$

where $\mathbf{x}$ is chosen to minimize the worst-case regularized residual in the uncertainty region. However, since the solution is found with respect to the worst possible perturbations on the observation matrix and the output vector in the uncertainty regions, it may be highly conservative [4, 5].

Here, we propose a novel regularized LS estimation approach that provides a tradeoff between performance and robustness in order to mitigate the conservative nature of the worst-case residual approach as well as to preserve robustness. The regret for not using the optimal regularized LS estimator is defined as the difference between the performances of an estimator and the optimal regularized LS estimator, i.e.,

$$\mathcal{R}(\mathbf{x}; \Delta\mathbf{H}, \Delta\mathbf{y}) \triangleq l(\mathbf{x}; \Delta\mathbf{H}, \Delta\mathbf{y}) - \min_{\mathbf{w}\in\mathbb{C}^n} l(\mathbf{w}; \Delta\mathbf{H}, \Delta\mathbf{y}). \tag{2}$$

By making such a regret definition, we force our estimator not to construct the symbol vector according to the worst possible scenario considering that it may be too conservative. Instead, we define the regret of any estimator by the difference in the estimation performances of that estimator and the optimal estimator knowing both observation matrix and output vector in hindsight, so that we achieve a tradeoff between robustness and estimation performance.

In the next section, we provide our approach to the described robust regularized LS estimation problems in detail.

## 3. ROBUST REGULARIZED LEAST SQUARES ESTIMATION METHOD

In this section, we provide a novel robust regularized LS estimator based on a certain minimax criterion. We search for the data vector $\mathbf{x}$, which is the solution to the following optimization problem

$$\min_{\mathbf{x}\in\mathbb{C}^n} \max_{||\Delta\mathbf{H}||\leq\delta_H, ||\Delta\mathbf{y}||\leq\delta_Y} \mathcal{R}(\mathbf{x}; \Delta\mathbf{H}, \Delta\mathbf{y}),$$

where $\mathcal{R}(\mathbf{x}; \Delta\mathbf{H}, \Delta\mathbf{y})$ is defined as in (2). Denoting $\tilde{\mathbf{H}} \triangleq \mathbf{H} + \Delta\mathbf{H}$, $\tilde{\mathbf{y}} \triangleq \mathbf{y} + \Delta\mathbf{y}$, we consider the second term of the regret in (2) and denote the optimal $\mathbf{w}$ solving the minimization operator as follows

$$\mathbf{w}^* \triangleq \arg\min_{\mathbf{w}\in\mathbb{C}^n} l(\mathbf{w}; \Delta\mathbf{H}, \Delta\mathbf{y}) = \left(\tilde{\mathbf{H}}^H\tilde{\mathbf{H}} + \mu\mathbf{I}\right)^{-1}\tilde{\mathbf{H}}^H\tilde{\mathbf{y}}.$$

Then, we have

$$l(\mathbf{w}^*; \Delta\mathbf{H}, \Delta\mathbf{y}) = \tilde{\mathbf{y}}^H\tilde{\mathbf{P}}^{-1}\tilde{\mathbf{y}},$$

where $\tilde{\mathbf{P}} \triangleq \mathbf{I}+\mu^{-1}\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H$ is guaranteed to be invertible since $\tilde{\mathbf{P}} > 0$. Considering the Taylor series expansion based on Wirtinger calculus [7] for $l(\mathbf{w}^*; \Delta\mathbf{H}, \Delta\mathbf{y})$ around $\tilde{\mathbf{H}} = \mathbf{H}$ and $\tilde{\mathbf{y}} = \mathbf{y}$ (i.e., around $\Delta\mathbf{H} = \mathbf{0}$ and $\Delta\mathbf{y} = \mathbf{0}$), we have

$$\begin{aligned} l(\mathbf{w}^*; \Delta\mathbf{H}, \Delta\mathbf{y}) &= l(\mathbf{w}^*; \mathbf{0}, \mathbf{0}) \\ &+ 2\,\mathrm{Re}\left\{\mathrm{Tr}\left(\nabla l(\mathbf{w}^*; \Delta\mathbf{H}, \Delta\mathbf{y})\big|^H_{\Delta\mathbf{H}=\mathbf{0},\Delta\mathbf{y}=\mathbf{0}}\big[\Delta\mathbf{H} \;\; \Delta\mathbf{y}\big]\right)\right\} \\ &+ O\left(|||[\Delta\mathbf{H} \;\; \Delta\mathbf{y}]|||^2\right). \end{aligned} \quad (3)$$

Note that in (3), the first order Taylor approximation is introduced in order to obtain a tractable solution. Clearly, the effect of using this approximation vanishes as $|||[\Delta\mathbf{H} \;\; \Delta\mathbf{y}]|||$ decreases and for distortions with larger $|||[\Delta\mathbf{H} \;\; \Delta\mathbf{y}]|||$, one can easily use higher order approximations instead. However, we observe through our simulations that even for relatively large perturbations, a satisfactory performance is obtained using this approximation.

Hence, after some algebra, we obtain

$$l(\mathbf{w}^*; \Delta\mathbf{H}, \Delta\mathbf{y}) \approx \kappa + \mathbf{d}^H\Delta\mathbf{h} + \Delta\mathbf{h}^H\mathbf{d} + \mathbf{b}^H\Delta\mathbf{y} + \Delta\mathbf{y}^H\mathbf{b},$$

where $\kappa \triangleq l(\mathbf{w}^*; \mathbf{0}, \mathbf{0})$, $\mathbf{d} \triangleq \mathrm{vec}(\mathbf{D}^H)$, $\Delta\mathbf{h} \triangleq \mathrm{vec}(\Delta\mathbf{H})$,

$$\begin{aligned} \mathbf{D} &\triangleq \frac{\partial l(\mathbf{w}^*; \Delta\mathbf{H}, \Delta\mathbf{y})}{\partial \tilde{\mathbf{H}}}\bigg|_{\Delta\mathbf{H}=\mathbf{0},\Delta\mathbf{y}=\mathbf{0}} \\ &= -\mathbf{P}^{-1}\mathbf{y}\mathbf{y}^H\mathbf{P}^{-1}\mathbf{H}, \end{aligned} \quad (4)$$

and

$$\begin{aligned} \mathbf{b} &\triangleq \frac{\partial l(\mathbf{w}^*; \Delta\mathbf{H}, \Delta\mathbf{y})}{\partial \tilde{\mathbf{y}}}\bigg|_{\Delta\mathbf{H}=\mathbf{0},\Delta\mathbf{y}=\mathbf{0}} \\ &= \mathbf{P}^{-1}\mathbf{y}, \end{aligned} \quad (5)$$

$\mathbf{P} \triangleq \mathbf{I} + \mu^{-1}\mathbf{H}\mathbf{H}^H$. Hence we can approximate the regret in (2) as follows

$$\begin{aligned} \mathcal{R}(\mathbf{x}; \Delta\mathbf{H}, \Delta\mathbf{y}) &\approx l(\mathbf{x}; \Delta\mathbf{H}, \Delta\mathbf{y}) \\ &- (\kappa + \mathbf{d}^H\Delta\mathbf{h} + \Delta\mathbf{h}^H\mathbf{d} + \mathbf{b}^H\Delta\mathbf{y} + \Delta\mathbf{y}^H\mathbf{b}). \end{aligned} \quad (6)$$

In the following theorem, we illustrate how the optimization problem in (6) can be put in an SDP form.

**Theorem:** *Let $\mathbf{H} \in \mathbb{C}^{m\times n}$ be a full rank matrix with $m \geq n$ and $\mathbf{y} \in \mathbb{C}^m$, both having deterministic additive perturbations $\Delta\mathbf{H} \leq \delta_H$ and $\Delta\mathbf{y} \leq \delta_Y$, respectively, then the approximated optimization problem*

$$\min_{\mathbf{x}\in\mathbb{C}^n} \max_{||\Delta\mathbf{H}||\leq\delta_H, ||\Delta\mathbf{y}||\leq\delta_Y} \mathcal{R}(\mathbf{x}; \Delta\mathbf{H}, \Delta\mathbf{y}), \quad (7)$$

*where $\mathcal{R}(\mathbf{x}; \Delta\mathbf{H}, \Delta\mathbf{y})$ is approximated as in* (6)*, is equivalent to solving the following SDP problem*

$$\min \gamma$$

*subject to*

$$\tau_1 \geq 0, \tau_2 \geq 0, \textit{ and}$$

$$\begin{bmatrix} \gamma+\kappa-\tau_1-\tau_2 & (\mathbf{y}-\mathbf{H}\mathbf{x})^H & \mathbf{x}^H & \delta_Y\mathbf{b}^H & \delta_H\mathbf{d}^H \\ \mathbf{y}-\mathbf{H}\mathbf{x} & \mathbf{I} & \mathbf{0} & -\delta_Y\mathbf{I} & \delta_H\mathbf{X} \\ \mathbf{x} & \mathbf{0} & \mu\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \delta_Y\mathbf{b} & -\delta_Y\mathbf{I} & \mathbf{0} & \tau_1\mathbf{I} & \mathbf{0} \\ \delta_H\mathbf{d} & \delta_H\mathbf{X}^H & \mathbf{0} & \mathbf{0} & \tau_2\mathbf{I} \end{bmatrix} \geq 0, \quad (8)$$

*according to the notation we derived, and where $\mathbf{X}$ is the $m \times mn$ matrix defined as $\mathbf{X} \triangleq \mathbf{x}^H \otimes \mathbf{I}$.*

*Proof of Theorem:* Considering the minimax problem defined in (7), we introduce the following reformulation

$$\min_{\mathbf{x}\in\mathbb{C}^n} \max_{||\Delta\mathbf{H}||\leq\delta_H, ||\Delta\mathbf{y}||\leq\delta_Y} \mathcal{R}(\mathbf{x}; \Delta\mathbf{H}, \Delta\mathbf{y}) = \min_{\mathbf{x}\in\mathbb{C}^n, \gamma\in\mathbb{C}} \gamma,$$

subject to

$$\mathcal{R}(\mathbf{x}; \Delta\mathbf{H}, \Delta\mathbf{y}) \leq \gamma, \; \forall\Delta\mathbf{H}, \Delta\mathbf{y} : ||\Delta\mathbf{H}|| \leq \delta_H, ||\Delta\mathbf{y}|| \leq \delta_Y, \quad (9)$$

where $\mathcal{R}(\mathbf{x}; \Delta\mathbf{H}, \Delta\mathbf{y})$ is defined as in (6). By applying the Schur complement to the constraints in (9), we compactly denote it in the matrix form as follows

$$\begin{bmatrix} \gamma+\kappa+\mathbf{d}^H\Delta\mathbf{h}+\Delta\mathbf{h}^H\mathbf{d}+\mathbf{b}^H\Delta\mathbf{y}+\Delta\mathbf{y}^H\mathbf{b} & (\tilde{\mathbf{y}}-\tilde{\mathbf{H}}\mathbf{x})^H & \mathbf{x}^H \\ \tilde{\mathbf{y}}-\tilde{\mathbf{H}}\mathbf{x} & \mathbf{I} & \mathbf{0} \\ \mathbf{x} & \mathbf{0} & \mu\mathbf{I} \end{bmatrix} \geq 0, \quad (10)$$

$\forall\Delta\mathbf{H}, \Delta\mathbf{y} : ||\Delta\mathbf{H}|| \leq \delta_H, ||\Delta\mathbf{y}|| \leq \delta_Y$. Rearranging terms in (10), we obtain

$$\begin{aligned} \begin{bmatrix} \gamma+\kappa & (\mathbf{y}-\mathbf{H}\mathbf{x})^H & \mathbf{x}^H \\ \mathbf{y}-\mathbf{H}\mathbf{x} & \mathbf{I} & \mathbf{0} \\ \mathbf{x} & \mathbf{0} & \mu\mathbf{I} \end{bmatrix} \geq \\ -\begin{bmatrix} \mathbf{d}^H \\ \mathbf{X} \\ \mathbf{0} \end{bmatrix} \Delta\mathbf{h} \begin{bmatrix} 1 & \mathbf{0} & \mathbf{0} \end{bmatrix} - \begin{bmatrix} 1 \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \Delta\mathbf{h}^H \begin{bmatrix} \mathbf{d} & \mathbf{X}^H & \mathbf{0} \end{bmatrix} \\ -\begin{bmatrix} \mathbf{b}^H \\ -\mathbf{I} \\ \mathbf{0} \end{bmatrix} \Delta\mathbf{y} \begin{bmatrix} 1 & \mathbf{0} & \mathbf{0} \end{bmatrix} - \begin{bmatrix} 1 \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \Delta\mathbf{y}^H \begin{bmatrix} \mathbf{b} & -\mathbf{I} & \mathbf{0} \end{bmatrix}, \end{aligned} \quad (11)$$

$\forall\Delta\mathbf{H}, \Delta\mathbf{y} : ||\Delta\mathbf{H}|| \leq \delta_H, ||\Delta\mathbf{y}|| \leq \delta_Y$, where we used $\Delta\mathbf{H}\mathbf{x} = \mathbf{X}\Delta\mathbf{h}$. By applying Proposition 2 of [4] twice to (11), it follows that (7) is equivalent to (8), hence the desired result. Therefore, this concludes the proof of the theorem. □

**Corollary:** The results of this theorem implies that one can straightforwardly consider the following two special cases of the theorem, in which *i)* the uncertainty is only on the observation matrix, *ii)* the uncertainty is only on the output vector. i.e., $\delta_Y = 0$ and $\delta_H = 0$ cases, respectively. For $\delta_Y = 0$ case, one can reduce the size of the SDP constraint matrix by simply omitting the fourth column and row of the (8) to reduce the computational complexity. In a similar fashion, for $\delta_H = 0$ case, fifth column and row of the (8) can be omitted to achieve a lower computational complexity.

**Remark:** In the proof of the theorem, we use Proposition 2 of [4] that relies on the lossless *S*-procedure. However, *S*-procedure is lossless with two constraints when the corresponding two quadratic (Hermitian) forms on the complex linear space [8]. However, classical *S*-procedure for quadratic forms is, in general, lossy with two constraints in the real case [9]. Hence, the theorem cannot be extended for real linear space. On the other hand, under the frameworks presented in the Corollary, one can safely extend the same conclusions for the real case also, since *S*-procedure is lossless for quadratic forms with one constraint both in complex and real spaces.

## 4. SIMULATIONS

We provide numerical examples in different scenarios in order to illustrate the merits of the proposed algorithms. In the first set of the experiments, we randomly generate a observation matrix of size $m \times n$, and an output vector of size $m \times 1$, which are normalized to have unit norms. Then, we randomly generate $1000$ random perturbations $\Delta\mathbf{H}$, $\Delta\mathbf{y}$, where $||\Delta\mathbf{H}|| \leq \delta_H$, $||\Delta\mathbf{y}|| \leq \delta_Y$, $m = 5$, $n = 3$, $\delta_H = \delta_Y = 0.6$, and the regularization parameter is set as $\mu = 1$. Here, we label the robust regularized LS estimator we present in the theorem as "R-RLS", the estimator employing robust regularized LS algorithm of [1] as "M-RLS", and finally the direct regularized LS estimator tuned to the estimates of the observation matrix and the output vector [2] as "D-RLS", where we directly minimize $l(\mathbf{x}; \mathbf{0}, \mathbf{0})$ over $\mathbf{x} \in \mathbb{C}^n$.

The largest error for this scenario is $1.3826$ for the D-RLS method, $1.1069$ for the R-RLS method, and $1.1225$ for the M-RLS method. We theoretically should observe that since the M-RLS algorithm minimizes the worst-case regularized residual, it yields the best worst-case performance among all algorithms for these simulations. However, even observing $1000$ independent perturbations, such a pessimistic scenario did not occur. Yet, we note that for the worst-case perturbation, we expect M-RLS algorithm to minimize the regularized residual. However, due to this highly conservative nature, the overall performance of the M-RLS algorithm is significantly inferior to the D-RLS and the R-RLS algorithms. The smallest error for this experiment is $0.3271$ for the D-RLS method, $0.4758$ for the R-RLS method, and $0.5489$ for the M-RLS method. As can also be observed from Fig. 1, R-RLS algorithm we present in this paper provides a practical tradeoff between the performance and robustness. Moreover, the average error for this experiment is $0.7724$ for the D-RLS method, $0.7589$ for the R-RLS method, and $0.8447$ for the M-RLS method. This shows that our algorithm not only provides a tradeoff between performance and robustness, but also provides a better average performance compared to its competitors.

## 5. CONCLUSION

In this paper, we presented a robust and efficient approach to the regularized LS estimation problem with bounded data uncertainties based on a novel regret formulation. We obtained the data vectors minimizing the worst-case regrets by solving the SDP problem presented in the theorem. Through our simulations, we observed that the proposed robust regularized LS method provides not only an efficient tradeoff between the performance and robustness, better than the best available alternatives in different signal processing applications, but also improves the average error performance.

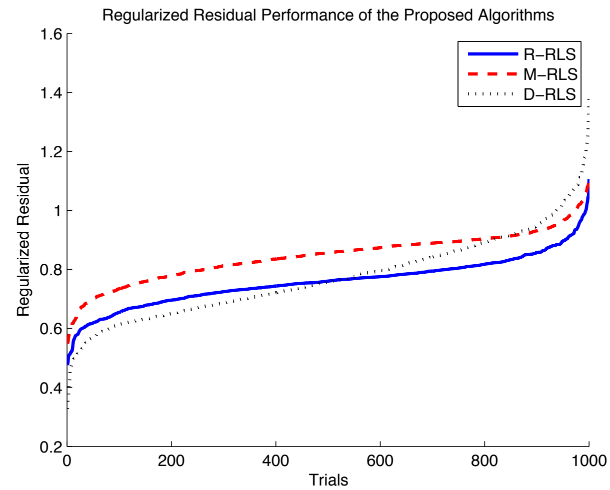


**Fig. 1:** Sorted regularized residuals for R-RLS, M-RLS, and D-RLS algorithms over 1000 trials when $m = 5$, $n = 3$, $\delta_H = \delta_Y = 0.6$, and $\mu = 1$.

## 6. REFERENCES


[1] L.E. Ghaoui and H. Lebret, "Robust solutions to least-squares problems with uncertain data," *SIAM J. Matrix Anal. Appl.*, vol. 18, no. 4, pp. 1035–1064, 1997.

[2] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice Hall, 2000.

[3] S.S. Kozat and A.T. Erdogan, "Competitive linear estimation under model uncertainties," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2388–2393, 2010.

[4] Y.C. Eldar and N. Merhav, "A competitive minimax approach to robust estimation of random parameters," *IEEE Trans. Signal Process.*, vol. 52, no. 7, pp. 1931–1946, 2004.

[5] Y.C. Eldar, A. Ben-Tal, and A. Nemirovski, "Linear minimax regret estimation of deterministic parameters with bounded data uncertainties," *IEEE Trans. Signal Process.*, vol. 52, no. 8, pp. 2177–2188, 2004.

[6] A.H. Sayed, V.H. Nascimento, and F.A.M. Cipparrone, "A regularized robust design criterion for uncertain data," *SIAM J. Matrix Anal. Appl.*, vol. 23, no. 4, pp. 1120–1142, 2001.

[7] A. Graham, *Kronecker Products and Matrix Calculus: with Applications*, John Wiley & Sons, 1981.

[8] A.L. Fradkov and V.A. Yakubovich, "The S-procedure and duality relations in nonconvex problems of quadratic programming," *Vestnik Leningradskogo Universiteta, Seriya Matematika*, , no. 1, pp. 101–109, 1973.

[9] A.L. Fradkov, "Duality theorems for certain nonconvex extremal problems," *Siberian Mathematical Journal*, vol. 14, no. 2, pp. 247–264, 1973.