Mathematics

http://www.aimspress.com/journal/Math

AIMS Mathematics, 11(6): 16095–16128.
DOI: 10.3934/math.2026662
Received: 07 April 2026
Revised: 08 May 2026
Accepted: 19 May 2026
Published: 05 June 2026

Research article

Novel robust logistic regression estimators for effectively modeling of multivariate binary data under outliers and multicollinearity: application to heavy metal contamination data in Al-Kharj landfills

Eslam Hussam[1], **Yousef Alharbi**[2], **Ahmed M. Gemeay**[3], **Samirah Alzubaidi**[4], **M. H. Harpy**[5], **Ramy Aldallal**[6], **M. S. Mohamed**[7] **and Ali T. Hammad**[3,*]

[1] Department of Accounting, College of Business Administration in Hawtat Bani Tamim, Prince Sattam bin Abdulaziz University, Saudi Arabia

[2] Department of Mathematics, College of Science, Qassim University, Buraydah 51452, Saudi Arabia

[3] Department of Mathematics, Faculty of Science, Tanta University, Tanta 31527, Egypt

[4] Department of Mathematics, Al-Qunfudah University College, Umm Al-Qura University, Mecca, Saudi Arabia

[5] Department of Mathematics, College of Science & Humanity Studies, Prince Sattam Bin Abdulaziz University, Saudi Arabia

[6] Department of Management, College of Business Administration in Hawtat Bani Tamim, Prince Sattam bin Abdulaziz University, Saudi Arabia

[7] Department of Mathematics, College of Science and Humanities in Hawtat Bani Tamim, Prince Sattam bin Abdulaziz University, Saudi Arabia

* **Correspondence:** Email: ali.taha@science.tanta.edu.eg.

Abstract: Logistic regression models are widely used for analyzing binary data, with the maximum likelihood estimator (MLE) being the standard method to estimate coefficients. However, the MLE becomes unstable and unreliable in the presence of multicollinearity or outliers. Outliers distort parameter estimates by unduly influencing the likelihood function, leading to bias and poor prediction. Multicollinearity inflates the variance of coefficients, reducing stability and interpretability. While biased estimators exist for multicollinearity and robust estimators for outliers, a unified framework that simultaneously handles both issues is still lacking. To address these issues, we have proposed a class of robust ridge-type estimators that combine robust logistic estimation with shrinkage methods. A comprehensive Monte Carlo simulation study evaluated the proposed estimators under varying levels of outliers and multicollinearity. Results show that our methods consistently outperform the traditional MLE and existing estimators in terms of accuracy and robustness. Finally, we demonstrated practical utility by analyzing heavy metal and metalloid contamination levels in landfill sites in Al-Kharj, Saudi Arabia, with empirical findings confirming that the proposed robust logistic estimators provide reliable and efficient inference when both multicollinearity and outliers are present.

Keywords: logistic regression model; biased estimator; robust estimator; multivariate heavy metal contamination data; outliers; multicollinearity
Mathematics Subject Classification: 62J07, 62J10, 62J12, 62P12

1. Introduction

The logistic regression model (LoRM) is one of the fundamental statistical methods for modeling dichotomous or binary outcomes [1]. The primary function of the LoRM is to describe the relationship between a binary response variable and one or more explanatory variables, and it is used in a multitude of applied fields such as biostatistics, finance, and even ecology. For instance, Jain and Srihari [2] applied the LoRM to compare the effectiveness of various machine learning algorithms in intrusion detection systems; Hocek et al. [3] used the LoRM in combination with machine learning for an exhaustive analysis of the probability of ship detention; Dey et al. [4] used the LoRM on data from a complex survey; and Hasim et al. [5] used the LoRM to analyze the effectiveness of agricultural extension services in clove farming.

The maximum likelihood estimator (MLE) is traditionally used to estimate the parameters of the LoRM. Nevertheless, this estimation procedure assumes that the explanatory variables are independent. In reality, this assumption is frequently disregarded due to multicollinearity, which renders the MLE statistically inefficient and leads to estimates of the coefficients that exhibit excessive variance and unstable signs, and complicate interpretation [6,7]. In an attempt to resolve this difficulty in the LoRM, numerous biased estimators have been suggested. In this regard, Schaefer et al. [8] proposed the ridge logistic estimator, Månsson et al. [9] proposed Liu estimators for the LoRM, Çetinkaya [10] proposed a jackknifed two-parameter estimator for the LoRM, Oladapo et al. [11] proposed a new two-parameter estimator for the LoRM, Lukman et al. [12] proposed the Kibria–Lukman estimator for the LoRM, and Awwad et al. [13] proposed a modified two-parameter estimator for the LoRM, among others.

The influence that outliers have on MLE performance causes estimation of parameters to become unstable and unreliable, which then leads to issues with the fit of the model as well as issues with the stability of the model. If observations are outliers or if observations deviate from the main and standard assumptions of the model, then robust estimation is paramount. A few robust estimators that are commonly used with linear models are Huber's M-estimator, the high-breakdown MM-estimator, the S-estimator with a tuned ρ-function, the least absolute deviation and least trimmed squares estimators, and many others [14]. In relation to the LoRM, a number of authors have proposed robust estimators that deal with outliers. Feng et al. [15] developed a robust logistic regression and classification framework within a neural information processing context. Building on earlier foundations, Bianco et al. [16] introduced robust estimation techniques specifically tailored to the LoRM. Carroll et al. [17] investigated the robustness properties of the LoRM from a statistical perspective. More recently, Hakimi et al. [18] proposed robust approaches for monitoring logistic regression profiles in the presence of outliers. Recent works provide several models that allow robust estimation of the LoRM. Some of these include robust functional logistic regression [19], penalized robust estimators for sparse settings [20], robust logistic regression with shift parameter estimation [21], and comparative studies of robust methods [22]. Moreover, comparative studies of neural networks provide additional explanations for the logistic regression determinants [23]. Collectively, these works signify the increasing attention toward the development of robustness for logistic regression in the context of outliers.

The presence of outliers and multicollinearity can negatively influence linear and generalized regression models. To deal with multicollinearity, the ridge estimator, Liu estimator, and other shrinkage approaches have been introduced, while M-estimation and MM-estimation have been suggested for outlier presence. Recent works have been able to tackle these concerns simultaneously. In particular, several researchers have merged these techniques to address such issues simultaneously within a single regression framework. Lukman et al. [24] proposed the robust Kibria–Lukman estimator for the LoRM, Arum [25] developed the robust Kibria–Lukman mixed with principal component estimator for linear regression, Mohammad et al. [26] introduced the robust two-parameter estimator for the Poisson regression, Hammad et al. [27] defined the robust ridge-type estimator for the beta regression, Alshangiti et al. [28] developed a robust ridge estimator for gamma regression, and Altukhaes et al. [29] proposed the robust Liu estimator for the restricted semiparametric regression. In the literature, however, there are not many studies that have examined optimal methods for the LoRM.

In this paper, we present new robust estimators for the LoRM that combine a novel biased estimator with effective methods such as the Bianco-Yuhai estimator and the conditionally unbiased bounded influence estimator. We combine these techniques to effectively address multicollinearity and outliers. Unlike other studies that treat this issue separately, our study stands out for its dual approach. Therefore, we model logistic regression in the presence of multicollinearity and outliers, using both an extended Monte Carlo simulation and its application to heavy metal contamination data in Al-Kharj landfills to evaluate the effectiveness and performance of the proposed estimator.

This paper will be organized in the following manner. We review the LoRM and the estimation methods in Section 2. In Section 3, the new estimator is presented along with its statistical characteristics, and a theoretical comparison is made to show how it is different from the other estimators. Section 4 is dedicated to the Monte Carlo simulation conducted to show the performance of the new estimator, which is the focus of Section 5 of the paper, where an example is given in numerical form. Section 6 will be devoted to the conclusions of the study.

2. Methodology

In this part, we explain the LoRM and the ordinary MLE. Then, we review existing robust and non-robust (biased) estimators to handle multicollinearity and outlier problems separately or simultaneously.

2.1. LoRM

The LoRM framework assumes each observation of the binary dependent variable y_i is an independent Bernoulli random variable, specifically $y_i \sim \text{Bernoulli}(\pi_i)$. The conditional probability

$\pi_i = P(y_i = 1|x_i)$ can be represented using the logistic function:

$$\pi_i = \frac{e^{x_i'\beta}}{1 + e^{x_i'\beta}}, \quad i = 1, 2, \ldots, n, \tag{2.1}$$

where x_i' represents the ith row of an $n\times(p+1)$ design matrix X and $\beta = (\beta_0, \beta_1, \ldots, \beta_p)'$ is the parameter vector.

The corresponding log-likelihood function, derived from the Bernoulli distribution, is expressed using the link function in Eq (2.1) as follows:

$$\ln L(\beta) = \sum_{i=1}^{n} \left[y_i \ln(\pi_i) + (1 - y_i) \ln(1 - \pi_i) \right]. \tag{2.2}$$

Maximizing the log-likelihood function in Eq (2.2) involves solving the score equations obtained by setting its first derivative to zero:

$$\frac{\partial \ln L(\beta)}{\partial \beta} = \sum_{i=1}^{n} (y_i - \pi_i) x_i' = 0. \tag{2.3}$$

Since Eq (2.3) is nonlinear, it is commonly solved using an iteratively reweighted least squares algorithm. At convergence, the MLE can be written in the closed form:

$$\hat{\beta}_{\text{MLE}} = (X'\hat{U}X)^{-1} X'\hat{U}\hat{z}, \tag{2.4}$$

where $\hat{U} = \text{diag}(\hat{u}_i)$ is a weight matrix with elements $\hat{u}_i = \hat{\pi}_i(1 - \hat{\pi}_i)$ and the components of the adjusted response vector $\hat{z}$ are given by:

$$\hat{z}_i = \ln(\hat{\pi}_i) + \frac{y_i - \hat{\pi}_i}{\hat{\pi}_i(1 - \hat{\pi}_i)}.$$

The MLE in the LoRM follows an asymptotic normal distribution. Its covariance matrix is given by the inverse of the expected information matrix, which is given as:

$$\left[-\mathrm{E} \left(\frac{\partial^2 \ln L(\beta)}{\partial \beta \partial \beta'} \right) \right]^{-1} = (X'\hat{U}X)^{-1}. \tag{2.5}$$

By employing the spectral decomposition of the estimated information matrix $X'\hat{U}X$, we consider an orthogonal matrix Q whose columns are the normalized eigenvectors of $X'\hat{U}X$ and λ_j are the eigenvalues of $X'\hat{U}X$. This matrix Q diagonalizes the information matrix such that $Q'(X'\hat{U}X)Q = \Lambda$, where $\Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{p+1})$ contains the corresponding eigenvalues arranged in non-increasing order. Hence,

$$\left[-\mathrm{E} \left(\frac{\partial^2 \ln L(\beta)}{\partial \beta \partial \beta'} \right) \right]^{-1} = Q\Lambda^{-1}Q', \tag{2.6}$$

$$\text{tr}\left[\left[-\mathrm{E} \left(\frac{\partial^2 \ln L(\beta)}{\partial \beta \partial \beta'} \right) \right]^{-1} \right] = \sum_{j=1}^{p+1} \frac{1}{\lambda_j}, \tag{2.7}$$

where tr[.] is the trace of the matrix. The harmful impact of multicollinearity on the MLE is mainly linked to the characteristics of the information matrix. When predictors show a high degree of correlation, the matrix $X'\hat{U}X$ becomes poorly conditioned, a situation indicated by the presence of small eigenvalues. This poor conditioning results in an increase in the asymptotic variance of the estimates, consequently compromising the stability of the regression coefficients.

2.2. *Robust logistic regression*

Several robust methods exist for the LoRM, some of which are implemented in statistical software. Two of such estimators are presented in the following:

2.2.1. Bianco and Yohai estimator

To enhance the robustness of the LoRM, Pregibon [30] proposed M-estimators that replace the standard deviation with a weighted function:

$$M(\beta) = \sum_{i=1}^{n} \rho\left(d^2(\pi_i(\beta), y_i)\right), \tag{2.8}$$

where $\rho(u)$ denotes a Huber-type loss function. Deviance residuals from this method measure the difference between what we actually see and what the model predicts, and their signs show whether the model is overestimating. Bianco and Yohai [16] pointed out an important problem: The original estimator did not reliably reduce the influence of high-leverage points and had some theoretical issues. Their refined estimator introduces an additive correction term:

$$M(\beta) = \sum_{i=1}^{n} \left[\rho\left(d^2(\pi_i(\beta), y_i)\right) + q(\pi_i(\beta))\right], \tag{2.9}$$

where $\rho(v)$ is a bounded, non-decreasing function, specified as:

$$\rho(v) = \begin{cases} v - \frac{v^2}{2k}, & \text{for } v \le k, \\ k/2, & \text{otherwise.} \end{cases} \tag{2.10}$$

This construction ensures both robustness against outliers and desirable theoretical properties.

2.2.2. Conditionally unbiased bounded influence estimator

Künsch et al. [31] developed a robust alternative by integrating M-estimation theory with a conditional unbiasedness constraint. Their estimator solves:

$$\sum_{i=1}^{n} \Psi(y_i, x_i, \beta) = 0, \tag{2.11}$$

where the score function Ψ is designed to satisfy $E(\Psi(y_i, x_i, \beta)|x_i) = 0$, ensuring Fisher consistency. The optimal form of this function is:

$$\Psi(y_i, x, \beta, b, B) = W(\cdot)\left\{y_i - g(\beta' x_i) - c\left(\beta' x_i, \frac{b}{h(x_i, B)}\right)\right\}. \tag{2.12}$$

A tuning parameter b constrains the estimator's sensitivity, while a dispersion matrix B and leverage measure $h(x, B)$ complete the framework. The function $c(\cdot)$ acts as a bias-correction term to preserve conditional unbiasedness. We define the corrected residual as:

$$r(y_i, x_i, \beta, b, B) = y_i - g(\beta' x_i) - c\left(\beta' x_i, \frac{b}{h(x_i, B)}\right). \tag{2.13}$$

The corresponding weights are computed as:

$$W(y_i, x_i, \beta, b, B) = W_b\left(r(y_i, x_i, \beta, b, B) \cdot h(x_i, B)\right), \tag{2.14}$$

where $W_b(x) = \min\left\{1, \frac{b}{|x|}\right\}$ is the Huber weight function. This Schweppe-type formulation systematically downweights observations exhibiting large products of corrected residuals and leverage. The dispersion matrix B is defined by the condition:

$$E\left(\Psi\Psi'\right) = B.$$

For detailed implementation, refer to Künsch et al. [31].

2.3. Existing estimators

Following the work of Hoerl and Kennard [32], Schaefer et al. [8] proposed the logistic ridge regression estimator (LRRE) as a more efficient estimator than the MLE for the LoRM under multicollinearity. The LRRE is defined as follows:

$$\hat{\beta}_{\text{LRRE}} = \left(X'\hat{U}X + k_{RR}I\right)^{-1} X'\hat{U}X\hat{\beta}_{\text{MLE}}, \quad k_{RR} > 0, \tag{2.15}$$

where k_{RR} is the parameter for the LRRE, I is the identity matrix of order $(p + 1) \times (p + 1)$, and if $k = 0$, then the LRRE simplifies to the MLE. The bias vector and the covariance matrix of the LRRE are given by:

$$\begin{aligned}
\text{Bias}\left(\hat{\beta}_{\text{LRRE}}\right) &= \text{E}\left(\hat{\beta}_{\text{LRRE}}\right) - \beta = -k_{RR}Q(\Lambda + k_{RR}I)^{-1}\alpha_{\text{ML}}, \\
\text{Cov}\left(\hat{\beta}_{\text{LRRE}}\right) &= Q(\Lambda + k_{RR}I)^{-1}\Lambda(\Lambda + k_{RR}I)^{-1}Q'.
\end{aligned}$$

The matrix mean squared error (MMSE) and mean squared error (MSE) for the LRRE can be calculated using the bias and covariance formulas as follows:

$$\begin{aligned}
\text{MMSE}\left(\hat{\beta}_{\text{LRRE}}\right) &= \text{Cov}\left(\hat{\beta}_{\text{LRRE}}\right) + \text{Bias}\left(\hat{\beta}_{\text{LRRE}}\right)\text{Bias}\left(\hat{\beta}_{\text{LRRE}}\right)' \\
&= Q(\Lambda + k_{RR}I)^{-1}\Lambda(\Lambda + k_{RR}I)^{-1}Q' + k_{RR}^2 Q(\Lambda + k_{RR}I)^{-1}\alpha_{\text{ML}}\alpha_{\text{ML}}'(\Lambda + k_{RR}I)^{-1}Q'
\end{aligned} \tag{2.16}$$

and

$$\text{MSE}\left(\hat{\beta}_{\text{LRRE}}\right) = \text{tr}\left[\text{MMSE}\left(\hat{\beta}_{\text{LRRE}}\right)\right] = \sum_{j=1}^{p+1} \frac{\lambda_j}{(\lambda_j + k_{RR})^2} + k_{RR}^2 \sum_{j=1}^{p+1} \frac{\alpha_{\text{ML}_j}^2}{(\lambda_j + k_{RR})^2}, \tag{2.17}$$

where $tr(.)$ is the trace of the matrix and $\alpha_{\text{ML}} = Q'\beta_{\text{MLE}}$.

Although the work of Schaefer et al. [8] made significant progress in addressing multicollinearity in the LoRMs, the LRRE was still susceptible to the presence of outliers in the data. Therefore, Lukman et al. [24] developed a ridge regression estimator within the realm of the LoRM by combining it with robust estimators such as the Bianco and Yohai estimator (BYE) and the conditionally unbiased bounded influence estimator (CE). This resulted in a robust biased estimator capable of simultaneously addressing multicollinearity and outliers. The LRRE with the BYE (LRR-BYE) is defined as follows:

$$\hat{\beta}_{\text{LRR-BYE}} = \left(X'\hat{U}X + k_{BY}I\right)^{-1} X'\hat{U}X\hat{\beta}_{\text{BYE}}, \quad k_{BY} > 0, \tag{2.18}$$

where $\hat{\beta}_{\text{BYE}}$ denotes the parameter estimates obtained using the BYE and k_{BY} refers to the robust shrinkage parameter for the LRR-BYE. Then the MSE and MMSE are given by:

$$\begin{aligned}\text{MMSE}\left(\hat{\beta}_{\text{LRR-BYE}}\right) &= Q(\Lambda + k_{BY}I)^{-1}\Lambda(\Lambda + k_{BY}I)^{-1}Q' \\ &+ k_{BY}^2 Q(\Lambda + k_{BY}I)^{-1}\alpha_{\text{BY}}\alpha'_{\text{BY}}(\Lambda + k_{BY}I)^{-1}Q'\end{aligned} \tag{2.19}$$

and

$$\text{MSE}\left(\hat{\beta}_{\text{LRR-BYE}}\right) = \text{tr}\left[\text{MMSE}\left(\hat{\beta}_{\text{LRR-BYE}}\right)\right] = \sum_{j=1}^{p+1}\frac{\lambda_j}{(\lambda_j + k_{BY})^2} + k_{BY}^2\sum_{j=1}^{p+1}\frac{\alpha_{\text{BY}_j}^2}{(\lambda_j + k_{BY})^2}, \tag{2.20}$$

where $\alpha_{\text{BY}} = Q'\beta_{\text{BYE}}$.

Similarly, Lukman et al. [24] proposed the LRR-CE, which combines the ridge regression estimator with the CE. The LRR-CE estimator of β is defined as:

$$\hat{\beta}_{\text{LRR-CE}} = \left(X'\hat{U}X + k_{CE}I\right)^{-1} X'\hat{U}X\hat{\beta}_{\text{CE}}, \quad k_{CE} > 0, \tag{2.21}$$

where $\hat{\beta}_{\text{CE}}$ denotes the parameter estimates obtained using the CE and k_{CE} refers to the robust shrinkage parameter for the LRR-CE. Then the MSE and MMSE are given by:

$$\begin{aligned}\text{MMSE}\left(\hat{\beta}_{\text{LRR-CE}}\right) &= Q(\Lambda + k_{CE}I)^{-1}\Lambda(\Lambda + k_{CE}I)^{-1}Q' \\ &+ k_{CE}^2 Q(\Lambda + k_{CE}I)^{-1}\alpha_{\text{CE}}\alpha'_{\text{CE}}(\Lambda + k_{CE}I)^{-1}Q'\end{aligned} \tag{2.22}$$

and

$$\text{MSE}\left(\hat{\beta}_{\text{LRR-CE}}\right) = \text{tr}\left[\text{MMSE}\left(\hat{\beta}_{\text{LRR-CE}}\right)\right] = \sum_{j=1}^{p+1}\frac{\lambda_j}{(\lambda_j + k_{CE})^2} + k_{CE}^2\sum_{j=1}^{p+1}\frac{\alpha_{\text{CE}_j}^2}{(\lambda_j + k_{CE})^2}, \tag{2.23}$$

where $\alpha_{\text{CE}} = Q'\beta_{\text{CE}}$.

3. Proposed robust and non-robust estimators

This section presents two main types of estimators for the LoRM. The first is the standard, non-robust estimator, which works well with multicollinear data but is distorted by outliers. Robust biased estimators, specifically designed to withstand outliers and multicollinearity, which yield more reliable results, comprise the second type. This section then theoretically compares the proposed estimator with existing estimators and concludes by discussing selection biasing parameters.

3.1. Non-robust logistic ridge-type estimator

By motivating the work done by Hoerl and Kennard [32] and Lukman et al. [33], we propose a new ridge-type estimator (RTE) by introducing a new formulation incorporating the constraint

$$\left(kd/k^{\frac{1}{2}}(1+d)^{\frac{1}{2}}\right)\hat{\beta} = k^{\frac{1}{2}}(1+d)^{\frac{1}{2}}\beta + \varepsilon'.$$

This augmentation strategy yields a biased but more stable estimator than LRRE, particularly effective in high-multicollinear scenarios.

The proposed estimator is derived through constrained optimization, minimizing the residual sum of squares $(y - X\beta)'(y - X\beta)$ subject to the quadratic constraint $(\beta - kd\hat{\beta})'(\beta - kd\hat{\beta}) = c$, by minimizing the following objective function:

$$(y - X\beta)'(y - X\beta) + k(1 + d)\left[(\beta - kd\hat{\beta})'(\beta - kd\hat{\beta}) - c\right].$$

With respect to β, this will yield the normal equation:

$$(X'\hat{U}X + k(1 + d)I_r)^{-1}\beta = X'y + kd\hat{\beta}, \tag{3.1}$$

where k is a non-negative constant. Solving Eq (3.1) yields the new estimator as follows:

$$\hat{\beta}_{\text{RTE}} = (X'\hat{U}X + k(1 + d)I)^{-1}(X'\hat{U}X + kd)\hat{\beta}, \quad 0 < d < 1,\ k > 0. \tag{3.2}$$

This new estimator is defined for the LoRM, named the new logistic ridge-type estimator (LRTE), and is defined as follows:

$$\hat{\beta}_{\text{LRTE}} = (X'\hat{U}X + k^*(1 + d^*)I)^{-1}(X'\hat{U}X + k^*d^*)\hat{\beta}_{\text{MLE}}, \quad 0 < d^* < 1,\ k^* > 0, \tag{3.3}$$

where k^* and d^* are the shrinkage parameters for LRTE. The proposed estimator includes several existing estimators as special cases. Specifically, if $k = 0$, then $\hat{\beta}_{\text{LRTE}} = \hat{\beta}_{\text{MLE}}$ and if $d = 0$, then $\hat{\beta}_{\text{LRTE}} = \hat{\beta}_{\text{LRRE}}$.

The bias vector and variance-covariance matrix of the proposed estimator are given by

$$\begin{aligned} \text{Bias}(\hat{\beta}_{\text{LRTE}}) = {} & \text{E}(\hat{\beta}_{\text{LRTE}}) - \beta = \left((\Lambda + k^*(1 + d^*)I)^{-1}(\Lambda + k^*d^*)\right)\beta - \beta \\ = {} & - k^*Q(\Lambda + k^*(1 + d^*)I)^{-1}\alpha_{\text{ML}} \end{aligned} \tag{3.4}$$

and

$$\begin{aligned} \text{Cov}(\hat{\beta}_{\text{LRTE}}) = {} & \text{E}\left(\left[\hat{\beta}_{\text{LRTE}} - \text{E}(\hat{\beta}_{\text{LRTE}})\right]\left[\hat{\beta}_{\text{LRTE}} - \text{E}(\hat{\beta}_{\text{LRTE}})\right]'\right) \\ = {} & \left[(\Lambda + k^*(1 + d^*)I)^{-1}(\Lambda + k^*d^*)\right]\text{Cov}(\hat{\beta}_{\text{MLE}})\left[(\Lambda + k^*(1 + d^*)I)^{-1}(\Lambda + k^*d^*)\right]' \\ = {} & Q(\Lambda + k^*(1 + d^*)I)^{-1}(\Lambda + k^*d^*)\Lambda(\Lambda + k^*(1 + d^*)I)^{-1}(\Lambda + k^*d^*)Q'. \end{aligned} \tag{3.5}$$

Using Eqs (3.4) and (3.5) to calculate the MMSE and MSE of the proposed estimator, we have

$$\begin{aligned} \text{MMSE}(\hat{\beta}_{\text{LRTE}}) = {} & \text{Cov}(\hat{\beta}_{\text{LRTE}}) + \text{Bias}(\hat{\beta}_{\text{LRTE}})\text{Bias}(\hat{\beta}_{\text{LRTE}})' \\ = {} & Q(\Lambda + k^*(1 + d^*)I)^{-1}(\Lambda + k^*d^*)\Lambda(\Lambda + k^*(1 + d^*)I)^{-1}(\Lambda + k^*d^*)Q' \\ & + k^{*2}Q(\Lambda + k^*(1 + d^*)I)^{-1}\alpha_{\text{ML}}\alpha'_{\text{ML}}(\Lambda + k^*(1 + d^*)I)^{-1}Q' \end{aligned} \tag{3.6}$$

and

$$\begin{aligned} \text{MSE}(\hat{\beta}_{\text{LRTE}}) &= \text{tr}\left(\text{MMSE}(\hat{\beta}_{\text{LRTE}})\right) \\ &= \sum_{j=1}^{p+1} \frac{(\lambda_j + k^*d^*)^2}{\lambda_j(\lambda_j + k^*(1 + d^*))^2} + k^{*2}\sum_{j=1}^{p+1} \frac{\alpha^2_{\text{ML}_j}}{(\lambda_j + k^*(1 + d^*))^2}. \end{aligned} \tag{3.7}$$

3.2. Robust logistic ridge-type estimators

The new proposed LRTE is effective in addressing multicollinearity in the LoRM; however, it is not sufficient to handle both multicollinearity and outliers. Therefore, we developed new estimators for the LoRM by combining the LRTE with two robust estimators: the BYE and the CE. These combinations, referred to as the LRT-BYE and LRT-CE, are designed to effectively manage both multicollinearity and outliers. The proposed LRT-BYE is defined as follows:

$$\hat{\beta}_{\text{LRT-BYE}} = (X'\hat{U}X + k^*_{BY}(1 + d^*_{BY})I)^{-1}(X'\hat{U}X + k^*_{BY}d^*_{BY})\hat{\beta}_{\text{BYE}}, \quad 0 < d^*_{BY} < 1,\ k^*_{BY} > 0, \tag{3.8}$$

where k^*_{BY} and d^*_{BY} are the shrinkage parameters for the LRT-BYE. The proposed LRT-BYE contains several existing estimators as special cases. Specifically, when (k^*_{BY}), the LRT-BYE reduces to the BYE, and when (d^*_{BY}), it simplifies to the LRR-BYE. Then the MSE and MMSE of LRT-BYE are given by:

$$\begin{aligned}\text{MMSE}(\hat{\beta}_{\text{LRT-BYE}}) = {} & Q(\Lambda + k^*_{BY}(1 + d^*_{BY})I)^{-1}(\Lambda + k^*_{BY}d^*_{BY})\Lambda(\Lambda + k^*_{BY}(1 + d^*_{BY})I)^{-1}(\Lambda + k^*_{BY}d^*_{BY})Q' \\ & + {k^*_{BY}}^2 Q(\Lambda + k^*_{BY}(1 + d^*_{BY})I)^{-1}\alpha_{\text{BY}}\alpha'_{\text{BY}}(\Lambda + k^*_{BY}(1 + d^*_{BY})I)^{-1}Q' \end{aligned} \tag{3.9}$$

and

$$\begin{aligned}\text{MSE}(\hat{\beta}_{\text{LRT-BYE}}) &= \operatorname{tr}\left(\text{MMSE}(\hat{\beta}_{\text{LRT-BYE}})\right) \\ &= \sum_{j=1}^{p+1} \frac{(\lambda_j + k^*_{BY}d^*_{BY})^2}{\lambda_j(\lambda_j + k^*_{BY}(1 + d^*_{BY}))^2} + {k^*_{BY}}^2 \sum_{j=1}^{p+1} \frac{\alpha^2_{\text{BY}_j}}{(\lambda_j + k^*_{BY}(1 + d^*_{BY}))^2}. \end{aligned} \tag{3.10}$$

Similarly, we propose the LRT-CE estimator, which combines the newly developed ridge-type regression estimator with the CE. The LRR-CE estimator of (β) is defined as follows:

$$\hat{\beta}_{\text{LRT-CE}} = (X'\hat{U}X + k^*_{CE}(1 + d^*_{CE})I)^{-1}(X'\hat{U}X + k^*_{CE}d^*_{CE})\hat{\beta}_{\text{CE}}, \quad 0 < d^*_{CE} < 1,\ k^*_{CE} > 0, \tag{3.11}$$

where k^*_{CE} and d^*_{CE} are the shrinkage parameters for the LRT-CE. The proposed LRT-CE encompasses several existing estimators as special cases. Specifically, when (k^*_{CE}), the LRT-CE reduces to the CE, and when (d^*_{CE}), it simplifies to the LRR-CE. Then the MSE and MMSE of the LRT-CE are given by the following:

$$\begin{aligned}\text{MMSE}(\hat{\beta}_{\text{LRT-CE}}) = {} & Q(\Lambda + k^*_{CE}(1 + d^*_{CE})I)^{-1}(\Lambda + k^*_{CE}d^*_{CE})\Lambda(\Lambda + k^*_{CE}(1 + d^*_{CE})I)^{-1}(\Lambda + k^*_{CE}d^*_{CE})Q' \\ & + {k^*_{CE}}^2 Q(\Lambda + k^*_{CE}(1 + d^*_{CE})I)^{-1}\alpha_{\text{CE}}\alpha'_{\text{CE}}(\Lambda + k^*_{CE}(1 + d^*_{CE})I)^{-1}Q' \end{aligned} \tag{3.12}$$

and

$$\begin{aligned}\text{MSE}(\hat{\beta}_{\text{LRT-CE}}) &= \operatorname{tr}\left(\text{MMSE}(\hat{\beta}_{\text{LRT-CE}})\right) \\ &= \sum_{j=1}^{p+1} \frac{(\lambda_j + k^*_{CE}d^*_{CE})^2}{\lambda_j(\lambda_j + k^*_{CE}(1 + d^*_{CE}))^2} + {k^*_{CE}}^2 \sum_{j=1}^{p+1} \frac{\alpha^2_{\text{CE}_j}}{(\lambda_j + k^*_{CE}(1 + d^*_{CE}))^2}. \end{aligned} \tag{3.13}$$

3.3. Theoretical comparisons between the estimators

3.3.1. Comparison between the LRT-CE and the MLE

The estimator $\hat{\beta}_{\text{LRT-CE}}$ performs better than the estimator $\hat{\beta}_{\text{MLE}}$ based on the MSE criterion. In other words, $\text{MSE}(\hat{\beta}_{\text{LRT-CE}}) - \text{MSE}(\hat{\beta}_{\text{MLE}}) < 0$ whenever $(\lambda_j + k^*_{CE}d^*_{CE})^2 + \lambda_j {k^*_{CE}}^2 \alpha^2_{\text{CE}_j} < \left(\lambda_j + k^*_{CE}(1 + d^*_{CE})\right)^2$, $k^*_{CE} > 0$, and $d^*_{CE} > 0$.

Proof. To compare the two estimators, consider the difference between their MSE values:

$$
\begin{aligned}
D_1 &= \text{MSE}(\hat{\beta}_{\text{LRT-CE}}) - \text{MSE}(\hat{\beta}_{\text{MLE}}) \\
&= \sum_{j=1}^{p+1} \frac{(\lambda_j + k^*_{CE}d^*_{CE})^2 + \lambda_j {k^*_{CE}}^2 \alpha^2_{\text{CE}_j}}{\lambda_j \left(\lambda_j + k^*_{CE}(1 + d^*_{CE})\right)^2} - \sum_{j=1}^{p+1} \frac{1}{\lambda_j} \\
&= \sum_{j=1}^{p+1} \frac{(\lambda_j + k^*_{CE}d^*_{CE})^2 + \lambda_j {k^*_{CE}}^2 \alpha^2_{\text{CE}_j} - \left(\lambda_j + k^*_{CE}(1 + d^*_{CE})\right)^2}{\lambda_j \left(\lambda_j + k^*_{CE}(1 + d^*_{CE})\right)^2}.
\end{aligned}
$$

The difference $D_1 < 0$ is satisfied when $(\lambda_j + k^*_{CE}d^*_{CE})^2 + \lambda_j {k^*_{CE}}^2 \alpha^2_{\text{CE}_j} < \left(\lambda_j + k^*_{CE}(1 + d^*_{CE})\right)^2$, which shows that $\hat{\beta}_{\text{LRT-CE}}$ is superior to $\hat{\beta}_{\text{MLE}}$ whenever $k^*_{CE} > 0$ and $d^*_{CE} > 0$ for all $(j = 1, 2, \ldots, p + 1)$. □

3.3.2. Comparison between the LRT-CE and the LRRE

The estimator $\hat{\beta}_{\text{LRT-CE}}$ performs better than the estimator $\hat{\beta}_{\text{LRRE}}$ based on the MSE criterion. In other words, $\text{MSE}(\hat{\beta}_{\text{LRT-CE}}) - \text{MSE}(\hat{\beta}_{\text{LRRE}}) < 0$ whenever

$$
(\lambda_j + k_{RR})^2 \left((\lambda_j + k^*_{CE}d^*_{CE})^2 + \lambda_j {k^*_{CE}}^2 \alpha^2_{\text{CE}_j}\right) < \lambda_j \left(\lambda_j + k^*_{CE}(1 + d^*_{CE})\right)^2 \left(\lambda_j + k^2_{RR}\alpha^2_{\text{ML}_j}\right),
$$

$k_{RR} > 0$, $k^*_{CE} > 0$, and $d^*_{CE} > 0$.

Proof. To compare the two estimators, consider the difference between their MSE values:

$$
\begin{aligned}
D_2 &= \text{MSE}(\hat{\beta}_{\text{LRT-CE}}) - \text{MSE}(\hat{\beta}_{\text{LRRE}}) \\
&= \sum_{j=1}^{p+1} \frac{(\lambda_j + k^*_{CE}d^*_{CE})^2 + \lambda_j {k^*_{CE}}^2 \alpha^2_{\text{CE}_j}}{\lambda_j \left(\lambda_j + k^*_{CE}(1 + d^*_{CE})\right)^2} - \sum_{j=1}^{p+1} \frac{\lambda_j + k^2_{RR}\alpha^2_{\text{ML}_j}}{(\lambda_j + k_{RR})^2} \\
&= \sum_{j=1}^{p+1} \frac{(\lambda_j + k_{RR})^2 \left((\lambda_j + k^*_{CE}d^*_{CE})^2 + \lambda_j {k^*_{CE}}^2 \alpha^2_{\text{CE}_j}\right) - \lambda_j \left(\lambda_j + k^*_{CE}(1 + d^*_{CE})\right)^2 \left(\lambda_j + k^2_{RR}\alpha^2_{\text{ML}_j}\right)}{\lambda_j \left(\lambda_j + k^*_{CE}(1 + d^*_{CE})\right)^2 (\lambda_j + k_{RR})^2}.
\end{aligned}
$$

The difference $D_2 < 0$ is satisfied when

$$
(\lambda_j + k_{RR})^2 \left((\lambda_j + k^*_{CE}d^*_{CE})^2 + \lambda_j {k^*_{CE}}^2 \alpha^2_{\text{CE}_j}\right) < \lambda_j \left(\lambda_j + k^*_{CE}(1 + d^*_{CE})\right)^2 \left(\lambda_j + k^2_{RR}\alpha^2_{\text{ML}_j}\right),
$$

which shows that $\hat{\beta}_{\text{LRT-CE}}$ is superior to $\hat{\beta}_{\text{LRRE}}$ whenever $k_{RR} > 0$, $k^*_{CE} > 0$, and $d^*_{CE} > 0$ for all $(j = 1, 2, \ldots, p + 1)$. □

3.3.3. Comparison between the LRT-CE and the LRR-BYE

The estimator $\hat{\beta}_{\text{LRT-CE}}$ performs better than the estimator $\hat{\beta}_{\text{LRR-BYE}}$ based on the MSE criterion. In other words, $\text{MSE}(\hat{\beta}_{\text{LRT-CE}}) - \text{MSE}(\hat{\beta}_{\text{LRR-BYE}}) < 0$ whenever

$$(\lambda_j + k_{BY})^2 \left((\lambda_j + k_{CE}^* d_{CE}^*)^2 + \lambda_j {k_{CE}^*}^2 \alpha_{\text{CE}_j}^2 \right) < \lambda_j \left(\lambda_j + k_{CE}^*(1 + d_{CE}^*) \right)^2 \left(\lambda_j + k_{BY}^2 \alpha_{\text{BY}_j}^2 \right),$$

$k_{BY} > 0$, $k_{CE}^* > 0$, and $d_{CE}^* > 0$.

Proof. To compare the two estimators, consider the difference between their MSE values:

$$\begin{aligned} D_3 &= \text{MSE}(\hat{\beta}_{\text{LRT-CE}}) - \text{MSE}(\hat{\beta}_{\text{LRR-BYE}}) \\ &= \sum_{j=1}^{p+1} \frac{(\lambda_j + k_{CE}^* d_{CE}^*)^2 + \lambda_j {k_{CE}^*}^2 \alpha_{\text{CE}_j}^2}{\lambda_j \left(\lambda_j + k_{CE}^*(1 + d_{CE}^*) \right)^2} - \sum_{j=1}^{p+1} \frac{\lambda_j + k_{BY}^2 \alpha_{\text{BY}_j}^2}{(\lambda_j + k_{BY})^2} \\ &= \sum_{j=1}^{p+1} \frac{(\lambda_j + k_{BY})^2 \left((\lambda_j + k_{CE}^* d_{CE}^*)^2 + \lambda_j {k_{CE}^*}^2 \alpha_{\text{CE}_j}^2 \right) - \lambda_j \left(\lambda_j + k_{CE}^*(1 + d_{CE}^*) \right)^2 \left(\lambda_j + k_{BY}^2 \alpha_{\text{BY}_j}^2 \right)}{\lambda_j \left(\lambda_j + k_{CE}^*(1 + d_{CE}^*) \right)^2 (\lambda_j + k_{BY})^2}. \end{aligned}$$

The difference $D_3 < 0$ is satisfied when

$$(\lambda_j + k_{BY})^2 \left((\lambda_j + k_{CE}^* d_{CE}^*)^2 + \lambda_j {k_{CE}^*}^2 \alpha_{\text{CE}_j}^2 \right) < \lambda_j \left(\lambda_j + k_{CE}^*(1 + d_{CE}^*) \right)^2 \left(\lambda_j + k_{BY}^2 \alpha_{\text{BY}_j}^2 \right),$$

which shows that $\hat{\beta}_{\text{LRT-CE}}$ is superior to $\hat{\beta}_{\text{LRR-BYE}}$ whenever $k_{BY} > 0$, $k_{CE}^* > 0$, and $d_{CE}^* > 0$ for all $(j = 1, 2, \ldots, p + 1)$. □

3.3.4. Comparison between the LRT-CE and the LRR-CE

The estimator $\hat{\beta}_{\text{LRT-CE}}$ performs better than the estimator $\hat{\beta}_{\text{LRR-CE}}$ based on the MSE criterion. In other words, $\text{MSE}(\hat{\beta}_{\text{LRT-CE}}) - \text{MSE}(\hat{\beta}_{\text{LRR-CE}}) < 0$ whenever

$$(\lambda_j + k_{CE})^2 \left((\lambda_j + k_{CE}^* d_{CE}^*)^2 + \lambda_j {k_{CE}^*}^2 \alpha_{\text{CE}_j}^2 \right) < \lambda_j \left(\lambda_j + k_{CE}^*(1 + d_{CE}^*) \right)^2 \left(\lambda_j + k_{CE}^2 \alpha_{\text{CE}_j}^2 \right),$$

$k_{CE} > 0$, $k_{CE}^* > 0$, and $d_{CE}^* > 0$.

Proof. To compare the two estimators, consider the difference between their MSE values:

$$\begin{aligned} D_4 &= \text{MSE}(\hat{\beta}_{\text{LRT-CE}}) - \text{MSE}(\hat{\beta}_{\text{LRR-CEE}}) \\ &= \sum_{j=1}^{p+1} \frac{(\lambda_j + k_{CE}^* d_{CE}^*)^2 + \lambda_j {k_{CE}^*}^2 \alpha_{\text{CE}_j}^2}{\lambda_j \left(\lambda_j + k_{CE}^*(1 + d_{CE}^*) \right)^2} - \sum_{j=1}^{p+1} \frac{\lambda_j + k_{CE}^2 \alpha_{\text{CE}_j}^2}{(\lambda_j + k_{CE})^2} \\ &= \sum_{j=1}^{p+1} \frac{(\lambda_j + k_{CE})^2 \left((\lambda_j + k_{CE}^* d_{CE}^*)^2 + \lambda_j {k_{CE}^*}^2 \alpha_{\text{CE}_j}^2 \right) - \lambda_j \left(\lambda_j + k_{CE}^*(1 + d_{CE}^*) \right)^2 \left(\lambda_j + k_{CE}^2 \alpha_{\text{CE}_j}^2 \right)}{\lambda_j \left(\lambda_j + k_{CE}^*(1 + d_{CE}^*) \right)^2 (\lambda_j + k_{CE})^2}. \end{aligned}$$

The difference $D_4 < 0$ is satisfied when

$$(\lambda_j + k_{CE})^2 \left((\lambda_j + k_{CE}^* d_{CE}^*)^2 + \lambda_j {k_{CE}^*}^2 \alpha_{\text{CE}_j}^2 \right) < \lambda_j \left(\lambda_j + k_{CE}^*(1 + d_{CE}^*) \right)^2 \left(\lambda_j + k_{CE}^2 \alpha_{\text{CE}_j}^2 \right),$$

which shows that $\hat{\beta}_{\text{LRT-CE}}$ is superior to $\hat{\beta}_{\text{LRR-CE}}$ whenever $k_{CE} > 0$, $k_{CE}^* > 0$, and $d_{CE}^* > 0$ for all $(j = 1, 2, \ldots, p + 1)$. □

3.4. Selection of biasing parameters

Several researchers, including Hoerl et al. [32], Kibria et al. [34], and Khalaf and Shukur [35], among others, have introduced various estimators for the parameters (k) and (d). The selection of the corresponding shrinkage parameters fundamentally determines the performance and reliability of these estimators. In particular, the values of (k) and (d) dictate the trade-off between bias and variance: Higher values tend to decrease variance at the cost of increased bias, whereas lower values have the opposite effect. Therefore, precise tuning of these parameters is crucial to achieving an optimal balance that enhances estimation accuracy, especially in the presence of severe multicollinearity.

The optimal values of the shrinkage parameters (k) and (d) are those that minimize the MSE of the proposed estimator, denoted by MSE($\hat{\beta}_{\text{LRTE}}$) in Eq (3.7). These optimal values can be expressed as follows:

$$k^* = \frac{\lambda_j}{\hat{\alpha}^2_{\text{ML}_j}\lambda_j - d}, \quad d^* = \frac{\lambda_j(k\hat{\alpha}^2_{\text{ML}_j} - 1)}{k}.$$

For the LRTE, we suggest the following parameters:

$$\hat{k}_1^* = (p+1)\min\left(\frac{\lambda_j}{\hat{\alpha}^2_{\text{ML}_j}\lambda_j - d^*}\right), \quad \hat{d}_1^* = \frac{1}{2}\min\left(\frac{\lambda_j(\hat{k}_1^*\hat{\alpha}^2_{\text{ML}_j} - 1)}{\hat{k}_1^*}\right), \tag{3.14}$$

$$\hat{k}_2^* = \text{median}\left(\frac{\lambda_j}{\hat{\alpha}^2_{\text{ML}_j}\lambda_j - d^*}\right), \quad \hat{d}_2^* = \frac{1}{2}\min\left(\frac{\lambda_j(\hat{k}_2^*\hat{\alpha}^2_{\text{ML}_j} - 1)}{\hat{k}_2^*}\right), \tag{3.15}$$

where the initial value of d^* equals

$$d^* = \min\left(\frac{\hat{\alpha}^2_{\text{ML}_j}}{\hat{\alpha}^2_{\text{ML}_j} + \frac{1}{\lambda_j}}\right).$$

For LRT-BYE and LRT-CE, we use the same parameters as in LRTE but with replace $\hat{\alpha}^2_{\text{ML}_j}$ by $\hat{\alpha}^2_{\text{BY}_j}$ and $\hat{\alpha}^2_{\text{CE}_j}$, respectively.

Building on the approach of Lukman et al. [24], the LRRE is specified using the following parameters:

$$\hat{k}_1 = \frac{1}{\sum_{j=1}^{p+1}\hat{\alpha}^2_{\text{ML}_j}}, \tag{3.16}$$

$$\hat{k}_2 = \frac{1}{\max\left(\hat{\alpha}^2_{\text{ML}_j}\right)}. \tag{3.17}$$

For the LRR-BYE and LRR-CE, we use the same parameters as in the LRRE but replace $\hat{\alpha}^2_{\text{ML}_j}$ by $\hat{\alpha}^2_{\text{BY}_j}$ and $\hat{\alpha}^2_{\text{CE}_j}$, respectively.

4. Monte Carlo simulation study

This section provides a comparative evaluation of various logistic regression estimators through a simulation approach. Simulation methodologies have been widely adopted in statistical literature to examine estimator performance in both generalized models and LoRMs.

We minimize the MSE, which depends on both $\hat{\beta}$ and β, under the restriction that $\beta'\beta = 1$. Following the framework established by Schaeffer et al. [8], we can develop the LoRM using principles analogous to those applied in linear regression. To generate multicollinear covariates, we implement the approach described by Hammad et al. [36]:

$$x_{ij} = \sqrt{(1-\rho^2)}\Omega_{ij} + \rho\Omega_{i(j+1)}, \quad i = 1, 2, \ldots, n, \ j = 1, 2, \ldots, p+1,$$

where Ω_{ij} represents independent standard normal random variables, while ρ quantifies the pairwise correlation among covariates. We examine four distinct correlation structures:

$$\rho = 0.8, 0.85, 0.90, 0.95, \quad \text{and} \quad 0.99.$$

The binary response variable is modeled using a Bernoulli distribution, $y_i \sim Be(\pi_i)$, with

$$\pi_i = \exp(x_i'\beta)/(1 + \exp(x_i'\beta)).$$

We consider six different sample sizes:

$$n = 50, 75, 100, 200, 300, \quad \text{and} \quad 400.$$

Three contamination conditions are investigated in this study. The first condition introduces 10% outliers in the covariates, and the second condition introduces 20% outliers in the covariates. All simulations are performed using RStudio. The outlier observations are randomly allocated to particular sample positions. Following Lukman [24], we generate the number of outliers at each designated position by multiplying the length of the response variable by 0.1 or 0.2, which is 10% or 20% of the total observations. We use the round() function in R to get the closest whole number for the number of outliers. We use the sample() function to randomly choose indices from the response variable y to add outliers. These indices correspond to the observations where we will add outliers. Also, following Olaluwoye et al. [37] and Hammad et al. [27, 38], outliers were incorporated into the first and third covariates (x_1 and x_3) by replacing their original values with data generated from a normal distribution with mean = 5 and standard deviation = 2, thereby creating outliers. We calculate the simulated MSE according to

$$\text{MSE}(\hat{\beta}) = \frac{1}{5000}\sum_{i=1}^{5000}\left(\hat{\beta}_i - \beta\right)^T\left(\hat{\beta}_i - \beta\right). \tag{4.1}$$

In our simulation design, $\hat{\beta}_i$ corresponds to the estimated coefficient vector from the i-th iteration, whereas β represents the true parameter vector.

Tables 1–8 and Figures 1–3 show the results of the simulations, describe the effect of multicollinearity, sample sizes, explanatory variables, and outliers, and visualize how they affect the performance of the estimators.

Table 1. Simulated MSE values of various estimators within the LoRM at $p = 3$ under 10% outliers.

ρ	n	Non-Robust			Robust				Non-Robust		Robust			
		MLE	LRRE		LRR-BYE		LRR-CE		LRTE		LRT-BYE		LRT-CE	
		-	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$
0.80	50	1.793	1.2851	1.21729	1.04223	1.04777	0.97973	1.1219	0.58194	**0.48845**	0.50054	0.51113	0.50353	0.50914
	75	1.08485	0.85566	0.80132	0.72754	0.70296	0.69651	0.74064	0.51803	0.44705	0.4226	0.42853	**0.42012**	0.42304
	100	0.59739	0.51245	0.48814	0.45602	0.44223	0.44207	0.45966	0.35599	0.31894	**0.29757**	0.30245	0.29955	0.29945
	200	0.42334	0.38063	0.36777	0.34904	0.34087	0.34154	0.3504	0.29974	0.27752	**0.25948**	0.26347	0.25994	0.26022
	300	0.31863	0.29518	0.29082	0.27994	0.27866	0.2756	0.28399	0.24807	0.23188	0.22168	0.22416	**0.22163**	0.22205
	400	0.29294	0.27525	0.27159	0.26365	0.26227	0.26016	0.26655	0.23916	0.22664	0.215	0.21717	**0.2149**	0.21517
0.85	50	2.21443	1.59469	1.52792	1.29618	1.3258	1.21788	1.41507	0.62645	**0.5376**	0.56096	0.58738	0.58313	0.5848
	75	1.19478	0.929	0.8678	0.78126	0.75618	0.74464	0.79997	0.51901	0.43075	0.42016	0.42755	**0.41751**	0.41993
	100	0.52443	0.46816	0.45305	0.42863	0.42031	0.41899	0.43234	0.36261	0.32967	0.3144	0.3184	**0.31321**	0.31373
	200	0.37427	0.34144	0.33399	0.31805	0.31477	0.31221	0.32203	0.27859	0.26044	**0.24976**	0.25283	0.24994	0.25077
	300	0.40246	0.36878	0.35914	0.34423	0.33842	0.33799	0.34626	0.29732	0.27816	**0.25825**	0.26193	0.26096	0.26057
	400	0.33644	0.31127	0.30656	0.29162	0.29047	0.2868	0.29626	0.25908	0.24525	**0.23006**	0.23389	0.23119	0.23129
0.90	50	2.12309	1.5002	1.40689	1.20037	1.1972	1.12324	1.2874	0.61787	**0.45636**	0.46456	0.4807	0.4618	0.46548
	75	1.40832	1.0431	0.97694	0.85014	0.83156	0.8026	0.88875	0.51664	0.42219	0.41035	0.42174	**0.40934**	0.41175
	100	0.73851	0.61403	0.5775	0.53315	0.51251	0.51317	0.53714	0.36647	0.33917	**0.3111**	0.31998	0.31909	0.31774
	200	0.61686	0.52208	0.49614	0.46097	0.44695	0.4451	0.46658	0.33842	0.2978	**0.27609**	0.28616	0.27967	0.279
	300	0.53656	0.46523	0.44645	0.4167	0.40649	0.40401	0.42217	0.31925	0.28483	**0.26158**	0.27013	0.26303	0.26254
	400	0.39773	0.35864	0.34855	0.32999	0.32463	0.32265	0.33367	0.27524	0.25544	0.23464	0.24065	0.23437	**0.23428**
0.95	50	4.5092	2.89134	2.86362	2.13068	2.33756	1.95907	2.53295	0.75566	0.52256	**0.50569**	0.53517	0.53135	0.53443
	75	3.27936	2.19173	2.06505	1.70696	1.73309	1.58331	1.87569	0.8003	0.47132	**0.44404**	0.46835	0.44922	0.45011
	100	1.49267	1.08259	0.9884	0.8679	0.82616	0.81443	0.89036	0.47762	0.3383	**0.3148**	0.33004	0.31646	0.31492
	200	1.17017	0.87748	0.79621	0.71326	0.67068	0.67357	0.71871	0.41696	0.31455	**0.28353**	0.30007	0.28456	0.2844
	300	0.86177	0.67953	0.6294	0.57262	0.54432	0.54499	0.57864	0.36166	0.29311	**0.26107**	0.27588	0.26366	0.26327
	400	0.76051	0.61002	0.56533	0.52047	0.49269	0.49681	0.52231	0.34309	0.27687	0.25261	0.26525	0.25308	**0.25206**
0.99	50	27.16239	15.76199	16.18781	11.76817	13.59859	10.75121	14.81198	3.48861	0.49314	0.50796	0.59858	**0.48777**	0.49072
	75	11.41979	6.48095	6.45643	4.63759	5.21361	4.21169	5.7294	1.7397	0.44279	0.42007	0.47305	**0.41777**	0.41933
	100	5.71873	3.31497	3.14162	2.39868	2.51025	2.18199	2.77243	1.0486	0.33659	0.30533	0.34092	0.3046	**0.30362**
	200	4.82768	3.03907	2.89421	2.26689	2.35619	2.08249	2.57092	0.93443	0.38186	**0.32944**	0.37282	0.33735	0.33498
	300	4.1465	2.56935	2.42726	1.90387	1.96553	1.74687	2.15026	0.81011	0.35486	**0.30087**	0.33919	0.30596	0.30286
	400	3.37861	2.14778	2.00627	1.61056	1.62823	1.48252	1.77887	0.66688	0.34126	**0.29255**	0.3285	0.29754	0.29472

Table 2. Simulated MSE values of various estimators within the LoRM at $p = 3$ under 20% outliers.

ρ	n	Non-Robust			Robust				Non-Robust		Robust			
		MLE	LRRE		LRR-BYE		LRR-CE		LRTE		LRT-BYE		LRT-CE	
		-	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$
0.80	50	2.86267	2.11419	1.98127	1.7092	1.68639	1.60503	1.80568	0.72963	0.62465	**0.61224**	0.63549	0.61348	0.61674
	75	1.60434	1.24905	1.15217	1.04022	0.98896	0.98741	1.05196	0.55527	0.50908	**0.48568**	0.49841	0.4867	0.4909
	100	0.78082	0.66943	0.62147	0.59466	0.55724	0.57471	0.58298	0.3931	0.35834	0.33177	0.34415	0.33029	**0.32876**
	200	0.6395	0.56109	0.52932	0.50695	0.48199	0.49243	0.5005	0.36372	0.3481	**0.3185**	0.32805	0.3202	0.3189
	300	0.51934	0.47271	0.45528	0.43252	0.42011	0.42325	0.4319	0.33881	0.325	**0.29443**	0.30181	0.29679	0.29628
	400	0.45751	0.42134	0.4098	0.39065	0.3834	0.38332	0.39252	0.3159	0.30504	**0.27985**	0.28759	0.28163	0.28163
0.85	50	2.80316	2.07874	1.92847	1.68197	1.64324	1.57888	1.76178	0.68324	0.59419	**0.57944**	0.62674	0.58922	0.59451
	75	1.52145	1.20775	1.11164	1.01982	0.96531	0.97224	1.02178	0.59372	0.53019	0.51638	0.53212	**0.51408**	0.51921
	100	0.84515	0.72912	0.68028	0.64986	0.61342	0.62898	0.63992	0.43555	0.40412	**0.37635**	0.38883	0.37777	0.3779
	200	0.66821	0.5885	0.55908	0.53264	0.51129	0.51782	0.53022	0.38572	0.36202	**0.33177**	0.34112	0.33244	0.33255
	300	0.52584	0.47554	0.45913	0.43829	0.42774	0.42842	0.43998	0.33142	0.3282	**0.29448**	0.30284	0.30135	0.30054
	400	0.49146	0.44783	0.43055	0.41625	0.40277	0.40741	0.41418	0.32051	0.30955	**0.28391**	0.29296	0.28773	0.28672
0.90	50	4.69765	3.23385	3.00823	2.53868	2.53538	2.35616	2.74383	0.85608	0.58043	0.58452	0.63173	**0.5731**	0.57393
	75	2.72548	2.03329	1.87463	1.65622	1.59595	1.55823	1.70949	0.68452	0.56814	**0.54527**	0.57535	0.54694	0.54945
	100	1.18236	0.96414	0.86569	0.83124	0.75462	0.7947	0.80102	0.45139	0.39552	0.37796	0.39771	0.37809	**0.37552**
	200	1.01485	0.84065	0.76474	0.7305	0.67082	0.70037	0.70953	0.41216	0.36635	0.3351	0.35341	0.33489	**0.33191**
	300	0.84117	0.72006	0.66927	0.64509	0.60414	0.62241	0.63315	0.38613	0.36239	**0.32398**	0.34088	0.32998	0.32822
	400	0.63263	0.55279	0.51987	0.49132	0.46629	0.4764	0.48549	0.34007	0.32067	0.28964	0.30332	0.28967	**0.28842**
0.95	50	8.6846	5.73855	5.44808	4.43084	4.54641	4.08563	4.9445	1.17077	0.56544	**0.55355**	0.6258	0.56278	0.5642
	75	5.30147	3.61022	3.31511	2.75104	2.6823	2.54859	2.91835	0.87427	0.51635	**0.51151**	0.56241	0.5141	0.51364
	100	1.86097	1.40599	1.2624	1.14795	1.06197	1.08093	1.14337	0.45419	0.39954	**0.34997**	0.37581	0.35644	0.35434
	200	1.61643	1.2312	1.0877	1.02366	0.92521	0.96662	0.99532	0.45454	0.37824	**0.33359**	0.35554	0.33751	0.33397
	300	1.37641	1.08635	0.96736	0.923	0.83156	0.87567	0.89158	0.41032	0.35161	0.32383	0.35433	0.32327	**0.3205**
	400	1.2742	1.01413	0.90657	0.86827	0.78498	0.82496	0.84024	0.39151	0.326	0.30234	0.33184	0.2978	**0.29431**
0.99	50	37.64933	23.01505	22.40442	16.80586	18.02289	15.34546	19.76417	4.18282	0.55577	0.52722	0.61678	**0.51005**	0.51392
	75	22.74909	14.24493	13.6907	10.54034	11.11923	9.64537	12.17167	2.62434	0.57285	0.5219	0.61339	0.51755	**0.51402**
	100	11.30118	7.44117	6.93311	5.69986	5.70464	5.26619	6.2159	1.56442	0.49453	0.44394	0.51974	0.44293	**0.43912**
	200	7.17879	4.67338	4.25232	3.54842	3.45284	3.26768	3.7829	0.84053	0.40042	**0.354**	0.4159	0.36208	0.3581
	300	6.09074	4.0274	3.65978	3.07079	2.97368	2.83519	3.25229	0.76058	0.39899	0.3487	0.4066	0.35313	**0.34745**
	400	6.08352	4.06635	3.67388	3.16738	3.05444	2.92971	3.33455	0.80172	0.40818	0.35345	0.41621	0.35662	**0.35008**

Table 3. Simulated MSE values of various estimators within the LoRM at $p = 5$ under 10% outliers.

		Non-Robust			Robust				Non-Robust		Robust			
ρ	n	MLE	LRRE		LRR-BYE		LRR-CE		LRTE		LRT-BYE		LRT-CE	
		-	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$
0.80	50	4.03455	3.02272	2.78401	2.48683	2.39771	2.33482	2.57358	0.7812	0.72231	**0.69962**	0.76803	0.72081	0.73142
	75	2.20814	1.74245	1.58411	1.48104	1.38461	1.40729	1.47259	0.66263	0.6054	0.58689	0.60602	**0.58196**	0.59294
	100	0.98306	0.8484	0.78316	0.76858	0.71436	0.74284	0.74791	0.44547	0.42428	**0.39582**	0.41016	0.39792	0.39746
	200	0.80594	0.71021	0.66354	0.64677	0.60735	0.62759	0.63249	0.38861	0.38027	**0.34759**	0.36145	0.35076	0.35019
	300	0.75132	0.67024	0.62634	0.61074	0.57168	0.59457	0.59331	0.40961	0.39715	0.35338	0.36649	0.35524	**0.35285**
	400	0.6149	0.55908	0.53097	0.51192	0.48803	0.50032	0.50345	0.36347	0.35699	**0.32055**	0.33299	0.3225	0.32226
0.85	50	4.8572	3.56839	3.3106	2.9139	2.8674	2.72826	3.07891	0.76133	0.70022	**0.65695**	0.76141	0.68986	0.69512
	75	2.74337	2.10126	1.90136	1.73644	1.62657	1.63939	1.73998	0.65643	0.5959	0.56044	0.59343	**0.55742**	0.55994
	100	1.07782	0.92606	0.85219	0.83039	0.76926	0.8017	0.80643	0.46273	0.44016	**0.40359**	0.41844	0.40802	0.40687
	200	0.8788	0.76888	0.71007	0.69667	0.64343	0.67526	0.6722	0.41861	0.40109	0.3712	0.38378	0.37221	**0.37033**
	300	0.6846	0.61579	0.58008	0.56457	0.53366	0.55064	0.55222	0.38476	0.37201	0.34214	0.35423	0.3424	**0.34125**
	400	0.64041	0.5781	0.54792	0.53249	0.50637	0.51963	0.52336	0.36211	0.35551	0.31932	0.33177	0.32096	**0.31887**
0.90	50	7.51086	5.32832	4.99189	4.14735	4.13456	3.85456	4.46944	0.92048	0.71241	**0.65525**	0.78786	0.70716	0.70825
	75	3.30256	2.49756	2.25988	2.07473	1.94399	1.9518	2.08861	0.60969	0.54752	**0.5096**	0.54684	0.51327	0.51557
	100	1.43132	1.1739	1.05122	1.01928	0.9177	0.97422	0.97516	0.47123	0.44073	**0.39796**	0.4156	0.4	0.39917
	200	1.3094	1.09191	0.98521	0.96065	0.87255	0.92078	0.92388	0.43601	0.42065	**0.37474**	0.3974	0.38085	0.37725
	300	0.96815	0.8431	0.7779	0.75649	0.69831	0.73168	0.73141	0.41259	0.39835	0.35372	0.37265	0.35518	**0.35172**
	400	0.87459	0.7607	0.70499	0.68317	0.63572	0.66109	0.66479	0.39652	0.3851	**0.33714**	0.35527	0.341	0.33774
0.95	50	10.89647	7.50078	6.9557	5.92263	5.82288	5.49273	6.32617	1.15361	0.70331	**0.67461**	0.77489	0.69396	0.70097
	75	8.29556	5.94608	5.52573	4.71599	4.64	4.39762	5.00319	0.93252	0.63331	**0.60811**	0.68938	0.60994	0.61203
	100	2.78357	2.12311	1.86115	1.76471	1.58728	1.66335	1.71006	0.52107	0.47392	**0.41394**	0.44903	0.41955	0.41688
	200	2.71678	2.08062	1.80977	1.73034	1.52967	1.63082	1.65198	0.49261	0.43131	0.37846	0.41999	0.38177	**0.37159**
	300	2.05412	1.61959	1.4137	1.36127	1.19936	1.29067	1.28778	0.4673	0.42479	0.3636	0.39775	0.36171	**0.35372**
	400	1.74176	1.40224	1.24291	1.19095	1.06574	1.13336	1.1384	0.43952	0.4163	0.34676	0.37905	0.35005	**0.34294**
0.99	50	53.95751	36.39052	35.7066	27.91719	29.93397	25.76932	32.34715	3.46053	0.65181	**0.56873**	0.67404	0.57954	0.58057
	75	33.41233	22.45281	21.43742	17.7318	18.19909	16.38866	19.75395	2.77227	0.60153	**0.53602**	0.64461	0.54	0.53807
	100	13.25407	9.13579	8.24233	7.26161	6.91097	6.74304	7.5276	1.2895	0.4997	0.44525	0.53503	0.44427	**0.44051**
	200	11.35308	7.96742	7.19183	6.35617	6.02107	5.91937	6.53307	1.17445	0.52041	0.44672	0.53849	0.44909	**0.43982**
	300	9.08605	6.40819	5.69976	5.05547	4.67654	4.70217	5.09822	0.86363	0.46833	0.39128	0.46719	0.39248	**0.38549**
	400	8.21748	5.84393	5.17273	4.67852	4.28969	4.36099	4.66819	0.81004	0.46351	0.39679	0.47675	0.38376	**0.37439**

Table 4. Simulated MSE values of various estimators within the LoRM at $p = 5$ under 20% outliers.

ρ	n	Non-Robust			Robust				Non-Robust		Robust			
		MLE	LRRE		LRR-BYE		LRR-CE		LRTE		LRT-BYE		LRT-CE	
		-	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$
0.80	50	7.24995	5.33225	5.08767	4.45785	4.52485	4.17216	4.84071	0.73304	0.93507	**0.65488**	1.04762	0.92865	0.93165
	75	3.30607	2.61305	2.38995	2.21204	2.06895	2.09727	2.20368	0.653	0.61167	0.58569	0.62844	**0.58511**	0.58794
	100	1.30997	1.12398	1.01576	1.00053	0.90076	0.96488	0.94819	0.50205	0.48324	0.4442	0.4589	0.44302	**0.4423**
	200	0.96666	0.85192	0.7839	0.77377	0.71073	0.75039	0.74249	0.42768	0.41927	**0.37668**	0.39003	0.38013	0.37983
	300	0.88014	0.78113	0.72168	0.70979	0.655	0.6896	0.68255	0.4097	0.4025	**0.36134**	0.37398	0.36431	0.36247
	400	0.80782	0.73061	0.68688	0.6752	0.63367	0.65877	0.65612	0.41777	0.41476	0.36763	0.38352	0.36936	**0.36761**
0.85	50	5.936	4.48653	4.21764	3.71702	3.67449	3.49192	3.92862	0.7703	0.80826	**0.70433**	0.89476	0.8021	0.81226
	75	2.93691	2.33157	2.12239	2.00153	1.86724	1.8992	1.98857	0.67492	0.64144	**0.61906**	0.6505	0.62235	0.63135
	100	1.4653	1.24207	1.11557	1.09452	0.98261	1.0528	1.03773	0.50869	0.49226	0.44404	0.45924	0.44452	**0.44353**
	200	1.09485	0.96229	0.8821	0.87083	0.79622	0.84435	0.83248	0.4823	0.47207	**0.42999**	0.44612	0.43289	0.43154
	300	1.02854	0.90619	0.82237	0.81912	0.73965	0.79436	0.77429	0.43966	0.42906	**0.39044**	0.40725	0.39304	0.39145
	400	0.7808	0.70292	0.65727	0.64561	0.60367	0.62918	0.62611	0.39233	0.38744	**0.34485**	0.35858	0.34688	0.34519
0.90	50	7.29757	5.42022	5.03238	4.43337	4.33063	4.15246	4.6531	0.84151	0.73142	0.72439	0.78393	**0.71321**	0.72576
	75	4.51001	3.46616	3.1333	2.87513	2.66544	2.71099	2.8611	0.73484	0.69074	**0.62791**	0.67755	0.63111	0.64348
	100	1.86557	1.53564	1.35451	1.33055	1.17461	1.27156	1.25097	0.50126	0.48508	**0.43765**	0.46309	0.44261	0.4388
	200	1.53758	1.29014	1.14091	1.14013	1.00364	1.09364	1.06592	0.46153	0.44523	0.39697	0.41938	0.39961	**0.39616**
	300	1.29087	1.10107	0.98316	0.97297	0.86288	0.93662	0.91266	0.43596	0.42336	0.36625	0.38339	0.36906	**0.36477**
	400	1.1462	0.99016	0.89473	0.89403	0.8055	0.86354	0.84705	0.43357	0.42509	0.37213	0.39185	0.37518	**0.36995**
0.95	50	19.34038	13.90973	13.29555	11.29022	11.46235	10.52809	12.32246	1.44383	0.82544	**0.75955**	0.94787	0.82319	0.82991
	75	9.5585	7.08562	6.5531	5.86096	5.71191	5.49162	6.13462	0.84301	0.69021	**0.67531**	0.7884	0.69352	0.6994
	100	3.97139	3.09512	2.7242	2.61553	2.35053	2.47385	2.52309	0.57337	0.52698	**0.45948**	0.50453	0.46457	0.46234
	200	3.36838	2.64833	2.31338	2.23446	1.98874	2.11463	2.13478	0.49914	0.48335	**0.41152**	0.45662	0.41755	0.41164
	300	2.30696	1.85863	1.61784	1.58146	1.38131	1.50438	1.48031	0.47793	0.45717	0.40129	0.43473	0.40384	**0.39465**
	400	2.22986	1.82057	1.57555	1.58473	1.36478	1.511	1.46259	0.46533	0.42849	0.37371	0.41075	0.37772	**0.36933**
0.99	50	99.36726	67.27411	63.97496	51.28606	52.309	47.36084	56.88813	6.53631	0.66958	0.64985	0.76418	**0.6437**	0.64598
	75	54.2387	37.69067	35.02948	29.29204	28.90681	27.16444	31.37604	3.81488	0.64451	0.63709	0.76308	0.62271	**0.61938**
	100	19.62914	14.29186	12.97944	11.71231	11.10718	10.95674	12.00138	1.49214	0.61498	**0.51512**	0.62223	0.52702	0.51978
	200	14.37585	10.56164	9.41878	8.62913	7.93565	8.08276	8.58845	1.04293	0.57123	0.47036	0.56893	0.47378	**0.46689**
	300	12.42085	9.1132	8.03931	7.36842	6.71119	6.8953	7.27735	0.9251	0.5257	0.43861	0.5271	0.45017	**0.4382**
	400	10.02946	7.44049	6.51183	6.10688	5.51576	5.72408	5.97548	0.75054	0.50956	0.4137	0.49019	0.42142	**0.41247**

Table 5. Simulated MSE values of various estimators within the LoRM at $p = 9$ under 10% outliers.

		Non-Robust			Robust				Non-Robust		Robust			
ρ	n	MLE	LRRE		LRR-BYE		LRR-CE		LRTE		LRT-BYE		LRT-CE	
		-	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$
0.80	50	12.06234	9.38441	9.09947	8.12458	8.3222	7.67736	8.80737	4.03208	3.96457	3.38205	3.7697	3.392	**3.17251**
	75	6.07672	4.88883	4.44108	4.20941	3.91799	4.00138	4.16641	0.84627	0.834	**0.74229**	0.81937	0.7429	0.74412
	100	2.33517	2.00542	1.75391	1.79648	1.5623	1.73116	1.65202	0.6293	0.61838	**0.55354**	0.57537	0.55453	0.55689
	200	1.92347	1.68279	1.47794	1.5332	1.33447	1.48327	1.40636	0.56105	0.55775	**0.48997**	0.51002	0.49075	0.49011
	300	1.59401	1.40892	1.24229	1.28616	1.11142	1.24717	1.16978	0.53788	0.53244	0.46849	0.48108	0.46838	**0.46504**
	400	1.32998	1.19658	1.07284	1.09447	0.96322	1.0656	1.00765	0.51085	0.51205	0.44644	0.46186	0.44748	**0.44409**
0.85	50	16.22035	12.23018	11.75732	10.05609	10.23144	9.43877	10.92079	7.03298	7.85185	5.9911	6.75963	5.90321	**5.78104**
	75	8.05043	6.37648	5.81758	5.41357	5.11338	5.12593	5.45149	1.74187	1.72366	0.82017	0.92343	0.81895	**0.81789**
	100	2.17529	1.88727	1.69102	1.73012	1.53438	1.67088	1.61595	0.60384	0.59971	**0.53264**	0.55482	0.53403	0.5348
	200	1.89641	1.64879	1.43974	1.49257	1.27925	1.44191	1.35213	0.54704	0.54064	0.49257	0.51048	0.49382	**0.49221**
	300	1.6295	1.442	1.27159	1.31811	1.14215	1.27839	1.20162	0.53636	0.53217	0.47132	0.49018	0.47233	**0.46833**
	400	1.37371	1.23651	1.1115	1.14252	1.01042	1.11239	1.05657	0.51222	0.51102	0.44866	0.46712	0.4504	**0.44535**
0.90	50	22.23669	16.6595	16.12559	13.55621	13.93781	12.70575	14.88799	9.02642	9.82605	6.15225	6.85516	6.13426	**6.11867**
	75	10.11199	8.02181	7.36067	6.98628	6.6046	6.62595	7.03045	2.86026	2.82079	0.88824	1.00173	**0.88447**	0.89259
	100	3.90942	3.2844	2.90435	2.96655	2.6294	2.84451	2.78943	0.63295	0.63122	**0.58084**	0.63384	0.58184	0.58584
	200	2.70559	2.30736	1.98909	2.07653	1.7603	1.99731	1.87196	0.55149	0.54662	0.49049	0.52339	0.49226	**0.48957**
	300	2.35013	2.03122	1.75811	1.81751	1.53517	1.75261	1.62934	0.53734	0.5387	0.46768	0.49595	0.4692	**0.46256**
	400	1.78511	1.567	1.37351	1.40963	1.21353	1.36431	1.28084	0.51245	0.51307	0.44719	0.46885	0.44786	**0.44275**
0.95	50	34.65539	26.10611	25.03829	22.10482	22.49999	20.76059	23.98812	12.48433	12.81939	1.78738	2.09284	1.77549	**1.75975**
	75	20.43404	15.95555	14.80438	13.91292	13.41587	13.15896	14.28494	5.89939	5.72455	**0.80809**	0.94718	0.80882	0.80959
	100	6.68064	5.47406	4.80294	4.78754	4.21131	4.56532	4.49534	0.6289	0.6144	**0.54291**	0.6059	0.54666	0.55145
	200	5.012	4.14941	3.62016	3.67474	3.20327	3.51133	3.41511	0.58216	0.58109	**0.48192**	0.53477	0.48287	0.48584
	300	4.27424	3.56243	3.03202	3.1378	2.63772	3.00232	2.82235	0.54765	0.54574	0.46721	0.51225	0.46895	**0.4611**
	400	3.75175	3.164	2.70945	2.82514	2.38139	2.70926	2.5417	0.53859	0.53593	0.45098	0.49372	0.45411	**0.44438**
0.99	50	150.3656	109.62927	104.47044	87.39523	88.08113	81.54541	94.79504	15.05216	13.74807	1.47678	1.71856	1.47976	**1.47106**
	75	95.40135	71.5402	65.52237	57.6863	54.69274	54.11148	58.91218	4.41257	0.75428	0.70377	0.82495	0.69279	**0.68883**
	100	34.45074	27.20122	24.01415	23.07369	20.69813	21.85217	22.20678	1.30629	0.64864	**0.55106**	0.66947	0.55757	0.55501
	200	26.15159	20.80783	18.48967	17.75975	16.00844	16.83888	17.13426	1.00076	0.64472	**0.51904**	0.63142	0.52106	0.5195
	300	21.59123	17.23938	15.09469	14.5766	12.7568	13.8266	13.71532	1.02503	0.64609	0.52787	0.63632	0.52723	**0.52005**
	400	18.02702	14.40702	12.29127	12.67644	10.7893	12.03683	11.62541	0.67865	0.56566	0.47392	0.57166	0.47279	**0.46461**

Table 6. Simulated MSE values of various estimators within the LoRM at $p = 9$ under 20% outliers.

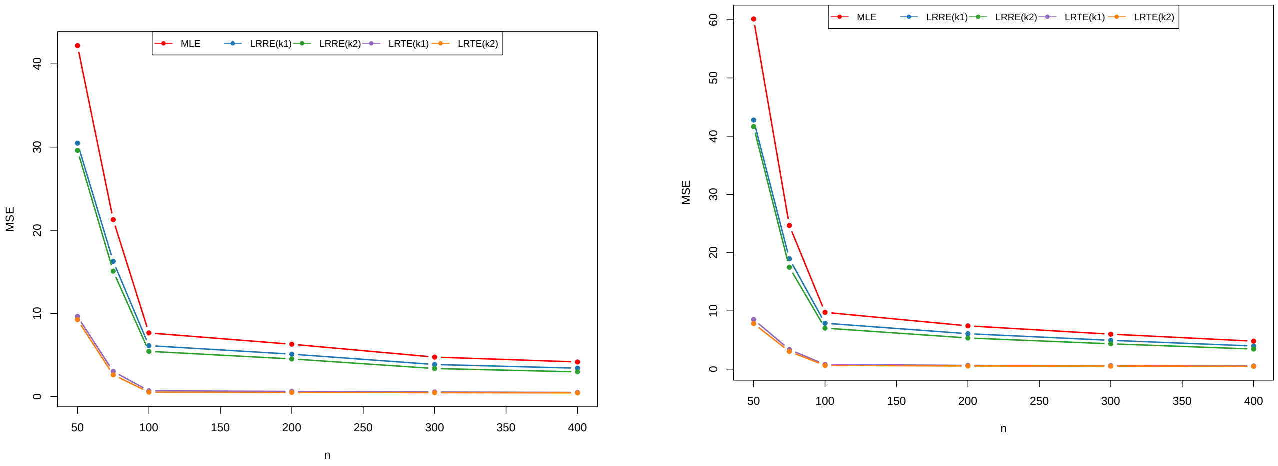
ρ	n	Non-Robust			Robust				Non-Robust		Robust			
		MLE	LRRE		LRR-BYE		LRR-CE		LRTE		LRT-BYE		LRT-CE	
		-	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$
0.80	50	14.1783	11.09245	10.68528	9.2971	9.4183	8.78436	9.98198	5.09224	5.00591	3.93105	4.25891	3.92758	**3.88632**
	75	8.25727	6.61938	6.01226	5.92903	5.51705	5.64259	5.8614	1.85662	1.84016	0.9549	1.06727	**0.95276**	0.95422
	100	2.73253	2.34553	2.05401	2.14835	1.85768	2.0703	1.96634	0.64775	0.64115	**0.58149**	0.60839	0.58247	0.58483
	200	2.06683	1.81551	1.59933	1.65578	1.43852	1.60334	1.51454	0.61502	0.6132	**0.54038**	0.55841	0.54118	0.54153
	300	1.92102	1.70984	1.5228	1.58466	1.39155	1.53911	1.45931	0.60213	0.60172	0.52306	0.54339	0.52469	**0.52278**
	400	1.51159	1.35849	1.20705	1.26308	1.10378	1.22927	1.15677	0.51622	0.5173	0.45648	0.47225	0.45738	**0.45456**
0.85	50	18.59826	14.02737	13.59779	11.51073	11.86207	10.80584	12.64317	8.07712	6.9402	8.62835	9.83003	8.41212	**7.87649**
	75	7.04178	5.73867	5.27998	5.02805	4.75244	4.79029	5.0322	2.83479	2.82426	**0.95164**	1.0539	0.95182	0.96359
	100	2.61108	2.26216	1.9943	2.08753	1.82047	2.01502	1.92262	0.62196	0.61496	**0.56899**	0.59423	0.56909	0.57455
	200	2.14046	1.87955	1.66194	1.72002	1.49996	1.66484	1.57941	0.59765	0.5926	**0.53281**	0.55385	0.53469	0.53354
	300	1.83463	1.63488	1.45105	1.51746	1.31958	1.47418	1.38579	0.56013	0.55987	0.50197	0.5211	0.5019	**0.50023**
	400	1.48227	1.33349	1.18612	1.23383	1.08019	1.20092	1.13148	0.51	0.511	0.44415	0.46018	0.4452	**0.44259**
0.90	50	25.271	19.22182	18.61759	15.76461	16.11431	14.81462	17.17258	10.13186	10.87561	4.49978	5.07323	4.48382	**4.16552**
	75	10.52682	8.47637	7.84257	7.32283	6.97346	6.95882	7.39552	2.84554	2.83418	1.00574	1.14053	**1.00448**	1.01325
	100	4.5469	3.8089	3.30304	3.42952	2.95497	3.28816	3.14364	0.68125	0.68288	0.61801	0.67601	0.62093	**0.61636**
	200	3.32808	2.86145	2.51651	2.59441	2.2642	2.49923	2.39475	0.58231	0.58678	**0.50491**	0.54066	0.50802	0.50957
	300	2.54783	2.21594	1.93768	2.00645	1.72498	1.93806	1.8236	0.59306	0.59549	**0.49785**	0.52142	0.49964	0.49858
	400	2.28177	2.00093	1.72357	1.83063	1.53243	1.77141	1.62355	0.53422	0.54082	0.47729	0.50286	0.4784	**0.47236**
0.95	50	38.38482	29.06824	27.64952	24.10817	24.26665	22.66026	25.90358	15.71827	15.97763	**5.2996**	6.28186	5.39543	5.3611
	75	21.90904	17.48227	16.28118	15.37599	14.79509	14.5967	15.69254	3.93272	3.80528	0.92534	1.08118	**0.91925**	0.92361
	100	7.867	6.50753	5.75853	5.89447	5.26687	5.63487	5.59701	0.66313	0.66482	**0.58513**	0.66165	0.5872	0.59105
	200	6.56817	5.48314	4.73066	4.87703	4.18272	4.66972	4.46129	0.63949	0.64008	**0.55327**	0.6187	0.55607	0.55424
	300	5.42801	4.57264	3.91786	4.02363	3.38087	3.85726	3.61003	0.61646	0.62075	0.52763	0.5838	0.53012	**0.52575**
	400	4.17528	3.5521	3.03356	3.23408	2.71951	3.10802	2.89841	0.54967	0.55503	0.47576	0.52168	0.4755	**0.47109**
0.99	50	233.43591	170.31258	161.73071	138.96812	140.59976	129.98448	150.71973	18.90251	17.7854	0.72767	0.81153	0.70804	**0.6977**
	75	94.79102	72.11489	65.48265	60.26814	56.74292	56.71232	60.95398	4.06208	3.76599	0.74799	0.87077	0.73033	**0.72731**
	100	42.14004	33.8058	30.35467	28.96231	26.23475	27.50443	28.03204	1.61862	0.72412	0.59823	0.72342	0.59886	**0.59464**
	200	32.87364	26.70103	23.53835	23.59741	20.80497	22.47103	22.25108	1.03589	0.64727	**0.55811**	0.68563	0.56469	0.55832
	300	25.89527	21.0635	18.44292	18.52832	16.15903	17.64831	17.29403	0.93599	0.65883	0.54189	0.65853	0.54215	**0.53658**
	400	18.69018	15.26459	13.17727	13.55388	11.69252	12.92162	12.52066	0.70987	0.61656	0.49558	0.59374	0.49748	**0.49303**

Table 7. Simulated MSE values of various estimators within the LoRM at $p = 12$ under 10% outliers.

		Non-Robust			Robust				Non-Robust		Robust			
ρ	n	MLE	LRRE		LRR-BYE		LRR-CE		LRTE		LRT-BYE		LRT-CE	
		-	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$
0.80	50	21.95336	17.04912	16.56775	14.93215	15.33363	14.12121	16.22344	6.27284	6.11071	5.23566	5.98123	5.22101	**4.78019**
	75	10.94924	8.9772	8.43523	8.19202	7.96711	7.82153	8.3906	3.93013	3.92341	**2.42514**	2.69024	2.4314	2.42665
	100	3.95229	3.41933	3.00878	3.05763	2.66367	2.94918	2.81408	0.7329	0.73276	0.65417	0.69712	**0.65413**	0.65704
	200	2.57948	2.28529	2.0048	2.11975	1.82542	2.05645	1.92147	0.70328	0.69964	**0.59517**	0.61997	0.59546	0.59645
	300	2.0645	1.84862	1.62957	1.7387	1.5017	1.69075	1.57697	0.60833	0.6048	**0.52255**	0.5404	0.52318	0.52508
	400	2.104	1.88871	1.65378	1.74232	1.48936	1.69502	1.56465	0.60756	0.60812	0.52763	0.54681	0.52792	**0.52511**
0.85	50	23.56389	17.94821	17.36018	14.93296	15.21889	14.06151	16.19584	8.10177	8.94544	6.17833	6.97103	5.81139	**5.74222**
	75	10.13595	8.33239	7.96372	7.41256	7.34469	7.07486	7.71837	3.90381	3.90506	**1.44206**	1.64425	1.44357	1.44737
	100	3.75061	3.26867	2.88826	3.00196	2.61929	2.90159	2.76157	0.73673	0.73582	0.65825	0.69885	**0.65804**	0.66596
	200	2.75637	2.43862	2.16497	2.24518	1.96369	2.17655	2.06381	0.66469	0.66508	0.57118	0.59683	0.57152	**0.57117**
	300	2.45637	2.18426	1.92196	2.01655	1.73742	1.95777	1.8272	0.65307	0.6525	**0.56453**	0.58889	0.56463	0.5652
	400	1.86169	1.68002	1.47506	1.56317	1.34041	1.52266	1.40654	0.59805	0.59513	0.5089	0.52321	0.50889	**0.50764**
0.90	50	43.97324	32.32382	31.56077	25.89112	26.85607	24.23212	28.73901	15.58779	14.00202	8.55198	9.70123	8.52905	**8.4495**
	75	17.88355	14.42629	13.47704	13.37432	13.03776	12.74345	13.75728	5.89918	5.84716	1.79819	2.10743	1.79409	**1.77626**
	100	5.47307	4.66626	4.16919	4.20424	3.75179	4.04223	3.96369	0.70785	0.71504	**0.63383**	0.69451	0.63497	0.64001
	200	4.32062	3.74379	3.26103	3.42423	2.9256	3.30456	3.09803	0.65795	0.66524	0.58646	0.636	**0.58623**	0.58909
	300	3.20996	2.81691	2.44399	2.62157	2.23018	2.53759	2.35653	0.6537	0.65543	0.55613	0.59014	0.55768	**0.55482**
	400	2.78674	2.47215	2.15712	2.269	1.94344	2.20057	2.0479	0.60715	0.61283	0.52354	0.55385	0.52393	**0.52026**
0.95	50	58.94148	44.24185	42.88317	36.57595	37.58517	34.36358	40.07067	21.74641	21.95881	11.11584	12.91442	11.09833	**9.85762**
	75	30.51095	24.57535	22.76726	22.24117	21.36315	21.17301	22.60406	7.05625	7.83258	1.47136	1.71397	**1.45847**	1.46554
	100	10.72461	9.05818	7.96455	8.1456	7.13773	7.81814	7.57453	0.77566	0.78891	**0.68056**	0.76444	0.682	0.68923
	200	8.19726	6.93652	5.97761	6.16885	5.22575	5.91837	5.57381	0.65437	0.66807	**0.57565**	0.64716	0.57713	0.57693
	300	6.41535	5.49335	4.76646	4.95817	4.22378	4.77072	4.48736	0.677	0.68485	**0.56928**	0.63086	0.57139	0.57077
	400	5.63521	4.85675	4.20131	4.39383	3.7235	4.23278	3.95501	0.60557	0.61215	0.52009	0.57578	0.52279	**0.51914**
0.99	50	341.06583	245.1097	240.35706	209.20451	220.92517	195.6197	236.11762	80.39071	80.0457	21.03429	20.13731	19.00994	**17.98323**
	75	147.71316	116.08108	107.2699	99.6467	94.41009	94.36323	100.64735	14.2816	13.8567	5.78888	4.90587	4.78402	**3.77296**
	100	51.11936	42.13	37.32297	37.67849	33.49578	35.99328	35.66383	1.52861	0.74369	**0.63863**	0.76725	0.64388	0.64241
	200	45.95693	38.358	33.98738	34.85908	30.74623	33.38786	32.6728	1.44425	0.73096	0.63231	0.77235	0.63201	**0.62916**
	300	29.65528	24.69239	21.18399	22.13365	18.73755	21.16845	20.06265	0.77698	0.61845	0.61336	0.61688	**0.51476**	0.51572
	400	27.49989	23.08239	19.87654	20.23852	17.05235	19.38171	18.24448	0.86324	0.70437	0.54854	0.65811	0.50997	**0.50842**

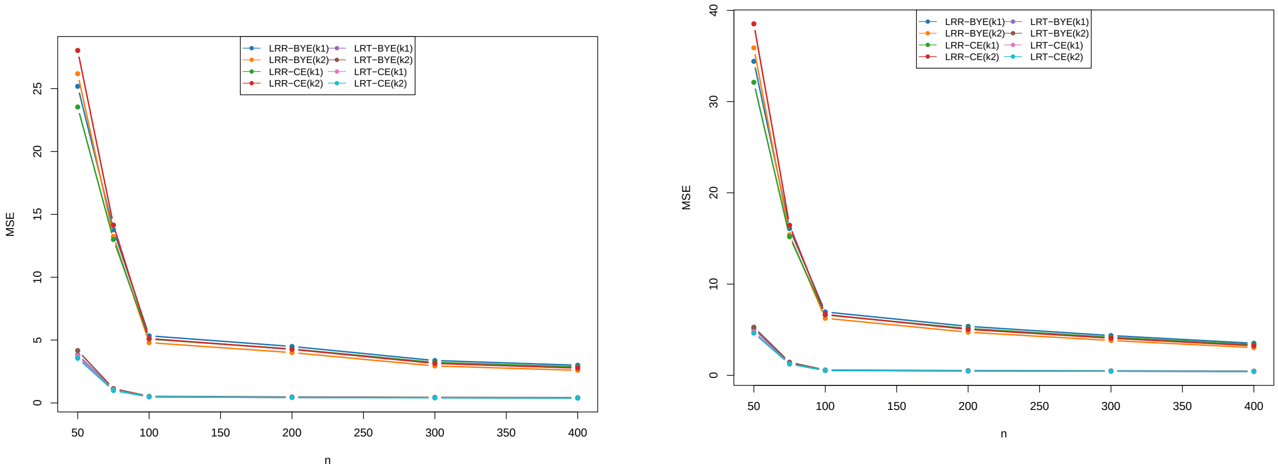
Table 8. Simulated MSE values of various estimators within the LoRM at $p = 12$ under 20% outliers.

		Non-Robust			Robust				Non-Robust		Robust			
ρ	n	MLE	LRRE		LRR-BYE		LRR-CE		LRTE		LRT-BYE		LRT-CE	
		-	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$
0.80	50	27.15136	20.86221	20.47288	17.84669	18.55121	16.84643	19.6473	13.30323	13.15061	8.74144	10.06726	8.66181	**8.4448**
	75	14.24761	11.6465	10.85697	10.40716	9.99026	9.92937	10.5472	5.95754	4.95498	1.70674	1.9652	1.71348	**1.69671**
	100	3.98839	3.45743	3.02385	3.27089	2.8391	3.1601	2.99663	0.75205	0.74847	**0.69544**	0.74313	0.69761	0.70309
	200	3.23909	2.85109	2.52094	2.60075	2.24972	2.51804	2.37063	0.68495	0.68501	**0.6062**	0.6428	0.60652	0.61029
	300	2.53419	2.26538	1.98271	2.09358	1.79186	2.03497	1.88389	0.67665	0.67297	0.58613	0.60855	0.58606	**0.58289**
	400	2.24309	2.02175	1.77198	1.88361	1.60778	1.83364	1.68945	0.58903	0.5889	0.51411	0.53195	0.5143	**0.51275**
0.85	50	27.95597	21.01423	20.50846	17.17945	17.81645	16.13118	18.98495	12.41396	11.08462	9.62448	10.87612	9.37042	**8.91434**
	75	11.94287	9.79365	9.20417	8.93803	8.65705	8.53672	9.11867	6.94833	6.94939	1.82262	2.05583	1.82385	**1.81049**
	100	4.47023	3.89708	3.51628	3.66979	3.29683	3.54714	3.4612	0.74547	0.75231	0.68939	0.74437	0.69066	**0.68872**
	200	2.87816	2.55146	2.22769	2.37721	2.03282	2.30648	2.14174	0.69508	0.68853	**0.58484**	0.60884	0.58548	0.59105
	300	2.82862	2.52258	2.19977	2.36364	2.00825	2.29563	2.11524	0.60723	0.61152	0.55859	0.59007	0.56003	**0.55818**
	400	2.14153	1.93832	1.71602	1.80958	1.56208	1.76341	1.63732	0.62106	0.62186	**0.53161**	0.54971	0.53311	0.53232
0.90	50	36.54802	27.55261	26.51137	22.87384	23.14941	21.50804	24.69295	15.69418	15.08558	12.19061	13.75471	12.16126	**11.90823**
	75	21.901	17.6543	16.62094	15.79344	15.39985	15.03262	16.27116	7.08939	7.97501	3.5461	4.07793	3.53846	**3.43654**
	100	6.51074	5.58827	4.98188	5.17429	4.63678	4.98436	4.88801	0.78414	0.7875	**0.72023**	0.78972	0.72264	0.72918
	200	4.56752	3.96541	3.45836	3.61626	3.09287	3.4916	3.27166	0.70667	0.71168	**0.60313**	0.65272	0.60375	0.60592
	300	3.69643	3.24102	2.79134	2.94622	2.4717	2.85067	2.61548	0.71277	0.71373	0.59229	0.62776	0.59149	**0.5902**
	400	3.31729	2.93773	2.53107	2.73404	2.30214	2.65137	2.4312	0.63225	0.63585	**0.55422**	0.58418	0.55573	0.55605
0.95	50	75.08921	56.20797	54.43837	47.12986	48.48256	44.32247	51.62022	21.34869	21.07524	15.62112	14.05344	13.75838	**13.51678**
	75	41.12243	33.09958	31.41044	30.00404	29.67899	28.56307	31.29852	9.20635	8.88863	2.78811	3.2289	2.77937	**2.75276**
	100	11.11823	9.41835	8.31059	8.7072	7.71058	8.36569	8.15813	0.73969	0.7477	0.67181	0.76283	**0.67127**	0.67545
	200	9.24184	7.9199	6.86367	7.24278	6.18174	6.97055	6.56575	0.63707	0.65736	**0.57586**	0.6524	0.57782	0.58343
	300	7.30181	6.28083	5.31522	5.69031	4.7111	5.48039	5.02016	0.6598	0.6624	**0.55002**	0.60996	0.55072	0.55439
	400	6.45208	5.57984	4.77912	5.12426	4.29734	4.9418	4.56593	0.63874	0.65292	**0.53999**	0.59764	0.54075	0.54074
0.99	50	510.09331	353.60513	352.26048	281.28054	302.85745	261.74736	325.46214	33.69677	33.05381	25.91925	22.97839	21.90683	**20.90413**
	75	153.49617	121.18729	110.44234	103.38293	96.87288	97.93014	103.39152	10.63121	10.88683	5.77698	5.87312	4.768	**4.75125**
	100	64.83269	54.004	47.96105	48.85603	43.26635	46.76701	45.99844	1.64168	1.74099	**0.65204**	0.78594	0.65357	0.65384
	200	49.20196	41.36841	36.13995	37.05554	32.26824	35.52639	34.34353	1.23546	0.75225	0.65811	0.79504	0.66002	**0.65805**
	300	38.77943	32.81565	28.94209	29.43716	25.72098	28.24838	27.32615	1.09397	0.72749	0.60244	0.72428	0.60406	**0.60173**
	400	29.84393	25.14737	21.57046	22.31823	18.83656	21.39659	20.12959	0.84675	0.68886	0.56419	0.66936	0.56421	**0.56328**



(a) Simulated MSE for the MLE and non-robust estimator in the LoRM under a 10% outlier with sample size.

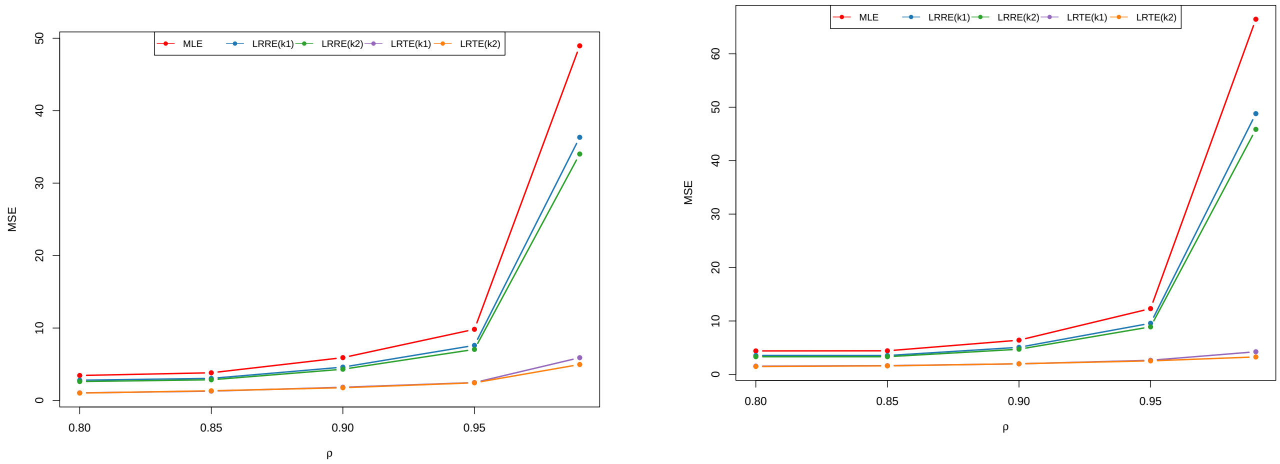
(b) Simulated MSE for the MLE and non-robust estimator in the LoRM under a 20% outlier with sample size.



(c) Simulated MSE for the robust estimator in the LoRM under a 10% outlier with sample size.

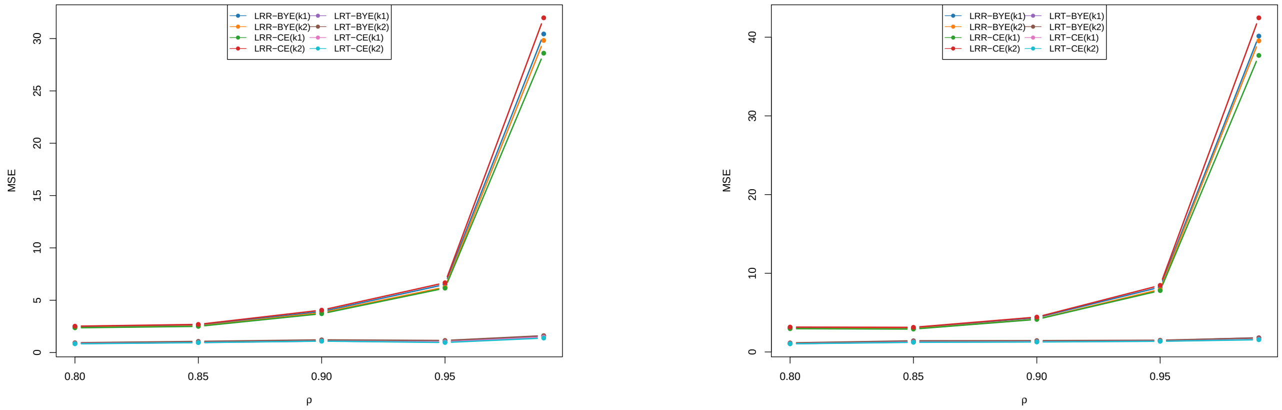
(d) Simulated MSE for the robust estimator in the LoRM under a 20% outlier with the sample size.

Figure 1. Simulated MSE for different estimators in the LoRM under outlier percentages with varying sample sizes.



(a) Simulated MSE for the MLE and non-robust estimator in the LoRM under a 10% outlier with multicollinearity levels.

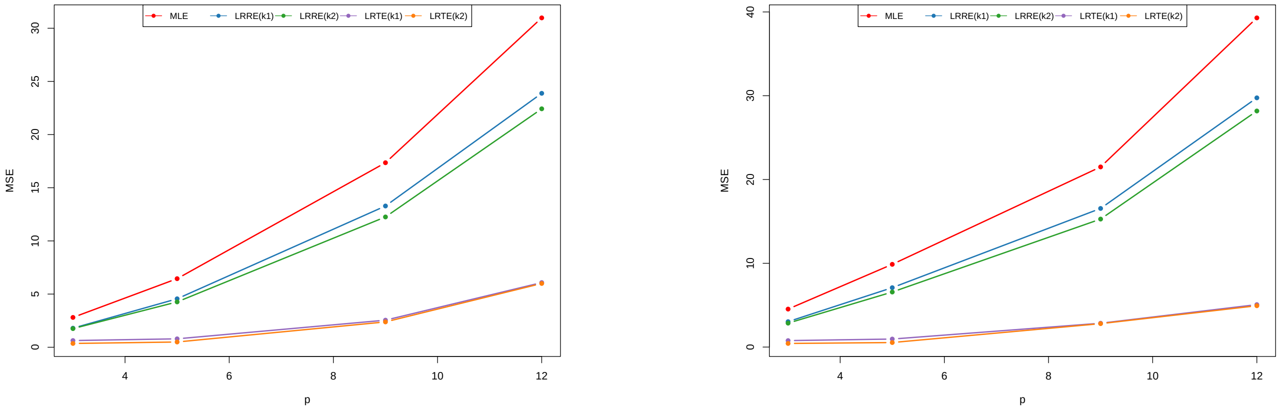
(b) Simulated MSE for the MLE and non-robust estimator in the LoRM under a 20% outlier with multicollinearity levels.



(c) Simulated MSE for the robust estimator in the LoRM under a 10% outlier with multicollinearity levels.

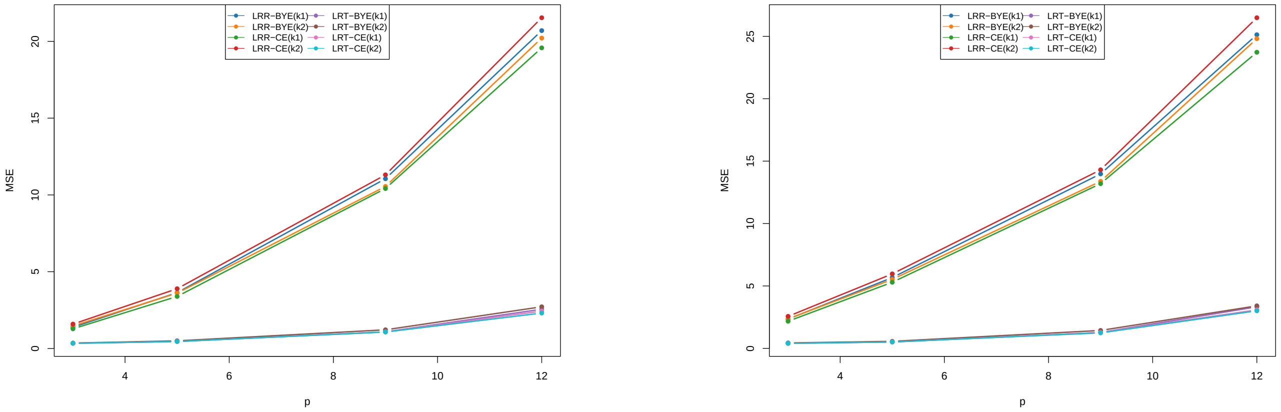
(d) Simulated MSE for the robust estimator in the LoRM under a 20% outlier with multicollinearity levels.

Figure 2. Simulated MSE for different estimators in the LoRM under outlier percentages with varying multicollinearity levels.



(a) Simulated MSE for the MLE and non-robust estimator in the LoRM under a 10% outlier with numbers of explanatory variables.

(b) Simulated MSE for the MLE and non-robust estimator in the LoRM under a 20% outlier with numbers of explanatory variables.



(c) Simulated MSE for the robust estimator in the LoRM under a 10% outlier with numbers of explanatory variables.

(d) Simulated MSE for the robust estimator in the LoRM under a 20% outlier with numbers of explanatory variables.

Figure 3. Simulated MSE for different estimators in the LoRM under outlier percentages with varying numbers of explanatory variables.

Tables 1–8 show the results of each of the simulations and the MSE for each of the logistic regression estimators for each of the testable estimators for each of the simulations for the LoRM and for the MSE for the non-robust estimators MLE, LRRE, and LRTE versus the estimators for the robust estimators LRR-BYE and LRR-CE versus the proposed robust estimators LRT-BYE and LRT-CE. Performance is evaluated at escalating degrees of multicollinearity (ρ = 0.80 to 0.99), varying sample sizes (n = 50 to 400), differing predictor dimensions (p = 3, 5, 9, 12), and distinct outlier contamination rates (10% and 20%). A consistent trend is evident, indicating that robust modified ridge-type estimators, especially those utilizing the second shrinkage method $(\hat{k}_2^*, \hat{d}_2^*)$, attain the minimal MSE in nearly all simulated scenarios.

In general, increasing the sample size (n) results in a decrease in the MSE for all estimators. For the MLE and LRRE estimators, increasing the sample size appears to have the largest effect on decreasing MSE, particularly in the presence of multicollinearity. For instance, in the robust MSE example in Table 1 ($\rho = 0.90$, $p = 3$) the MLE's MSE is 2.123 ($n = 50$) and 0.398 ($n = 400$). The robust estimators improved in efficiency with increasing sample size, but they still improved with small sample sizes (n), which helps to indicate that they are both robust and efficient.

Non-robust estimators perform worse when predictors have high correlation (ρ). As ρ approaches 0.99, the MSE for the MLE and standard LRRE increases significantly, and in some cases, the MSE is increased by orders of magnitude (in Table 7, $p = 12$, the MSE for the MLE is 341.07). The LRTE variants of the robust estimators, on the other hand, demonstrated remarkable strength. Even if ρ increased, the LRTE variants of the robust estimators' MSE did not increase a lot. This shows that the robust estimators are effective in controlling bias and variance, which is a defense against multicollinearity.

The effects of the curse of dimensionality are reflected in the increase in MSE across all methods with the increase in the number of predictors ($p = 3$ to $p = 12$). Despite this increase, the MSE results continue to support the same pattern of rank ordering. MSE results for the LRT-BYE and LRT-CE are the lowest, followed by the non-robust LRRE. This is the case for both levels of contamination, further indicating that the suggested robust methods can be paired with more sophisticated models.

With the addition of 10% and 20% outliers, non-robust estimator methods (MLE, LRRE, and LRTE) have a higher MSE, especially in high values for ρ and p. For instance, the MLE MSE exceeds 500 under 20% outliers, $p = 12$, and $\rho = 0.99$ (Table 8). Strong estimators avoid contamination very well. The LRT-BYE and LRT-CE, particularly, are able to provide accurate and stable estimates, demonstrating strong resilience to both vertical outliers and leverage points.

MLE: This method is extremely poor in performance because it has very high sensitivity to multicollinearity, outliers, and high dimensionality. It is therefore unwise to apply it in the presence of the data anomalies under consideration.

LRRE: This is better than the MLE if one only considers its performance on clean datasets; however, it is very sensitive to outliers.

LRR-BYE and LRR-CE: For outliers and multicollinearity, as well as for MSE reduction compared to their non-robust counterparts, these represent a big improvement. They are a decent consideration; however, classes of the LRTE are typically better.

LRT-BYE and LRT-CE: They are the best classes. The best bias-variance trade-off they achieve by using, in combination, the shrinkage of the LRTE and strong weighting. They have the least MSE; thus, they are the most efficient and most stable in environments characterized by multicollinearity, outlier contamination, small sample size, high dimension, etc. The version utilizing the second set of tuning parameters ($\hat{k}_2^*, \hat{d}_2^*$) always yields the best results.

5. Application

The performance of the proposed robust estimator, together with several existing estimators, was evaluated using the heavy metal contamination dataset from Al-Kharj landfills, originally introduced by

Alharbi et al. [39]. The dataset comprises 32 soil samples collected from two landfill sites in Al-Kharj, with the binary response variable (y_i) coded as 1 for samples originating from the first landfill and 0 for those from the second landfill. Each sample was analyzed for a set of heavy metals and metalloids using established contamination indices and soil quality standards. In this study, the explanatory variables consist of the measured concentrations of nine heavy metals: arsenic (As), cobalt (Co), chromium (Cr), copper (Cu), manganese (Mn), nickel (Ni), lead (Pb), vanadium (V), and zinc (Zn), respectively, denoted as (x_1, x_2, $\ldots$, x_9). These variables serve as fundamental predictors that enable a rigorous evaluation of the proposed estimator's capability to accurately and robustly distinguish between the two landfill sites, even in the presence of data contamination, high multicollinearity, and potential model instability.

To assess the presence of multicollinearity in the data, we use three basic measures: the correlation matrix, the conditional number, and the variance inflation factor. Figure 4 shows the correlation matrix for the explanatory variables, where high correlations are observed between the variables. The condition number is the square root of the ratio between the largest and smallest eigenvalues, which equals 744.0842. We also find that the variance inflation factor gives the following values:

$$5.20, 32.78, 15.39, 11.53, 8.94, 14.91, 4.15, 10.27, \text{ and } 12.19.$$

We find that all three measures indicate a high degree of multicollinearity among the explanatory variables, which affects the reliability of the parameters.

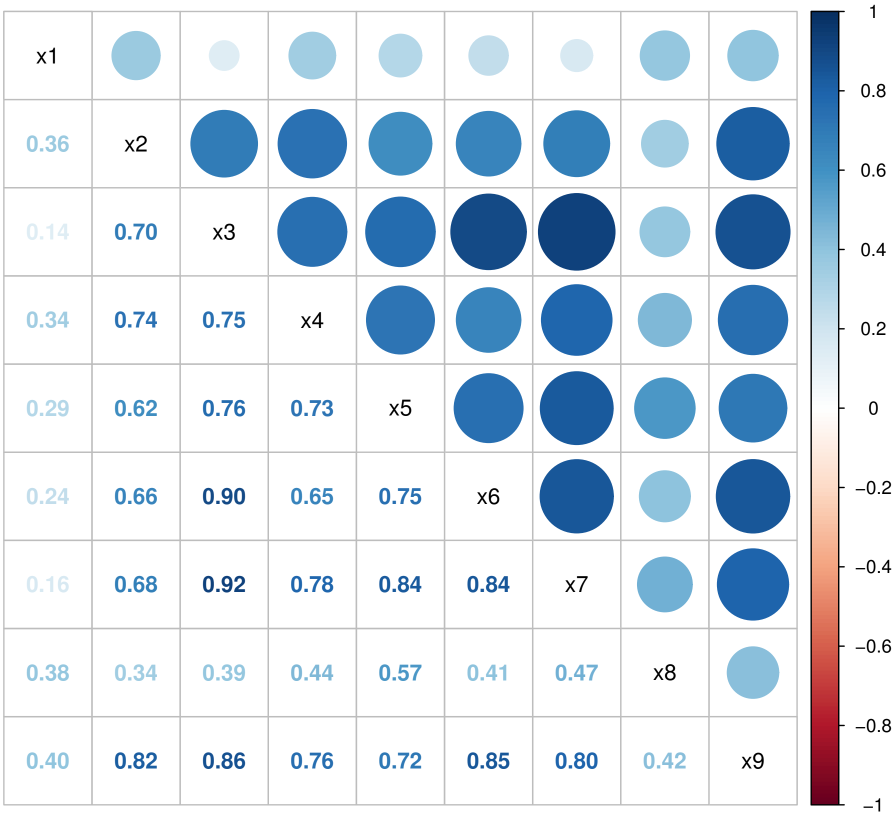


Figure 4. Pearson's correlation matrix of the explanatory variables in the data set.

The diagnostic plots for the LoRM in Figure 5 indicate the presence of a few outlying observations. While most residuals cluster around zero, several extreme residuals exceed ±2, suggesting potential outliers. The normal Q-Q plot confirms that the majority of residuals follow the expected distribution, with deviations only at the tails. The residuals-versus-fitted-values plot and Cook's distance analysis further highlight these influential points, which surpass the conventional $4/n$ threshold and may disproportionately impact the model estimates. Although the overall fit appears reasonable, these outliers and influential observations should be carefully examined to ensure the robustness of the model.

The boxplots for y and x_1– x_9 in Figure 6 show several outliers, particularly in x_2, x_3, x_4, x_6, x_7, and x_9. While most values fall within a typical range, these extreme observations may influence model estimates and should be considered in the analysis.

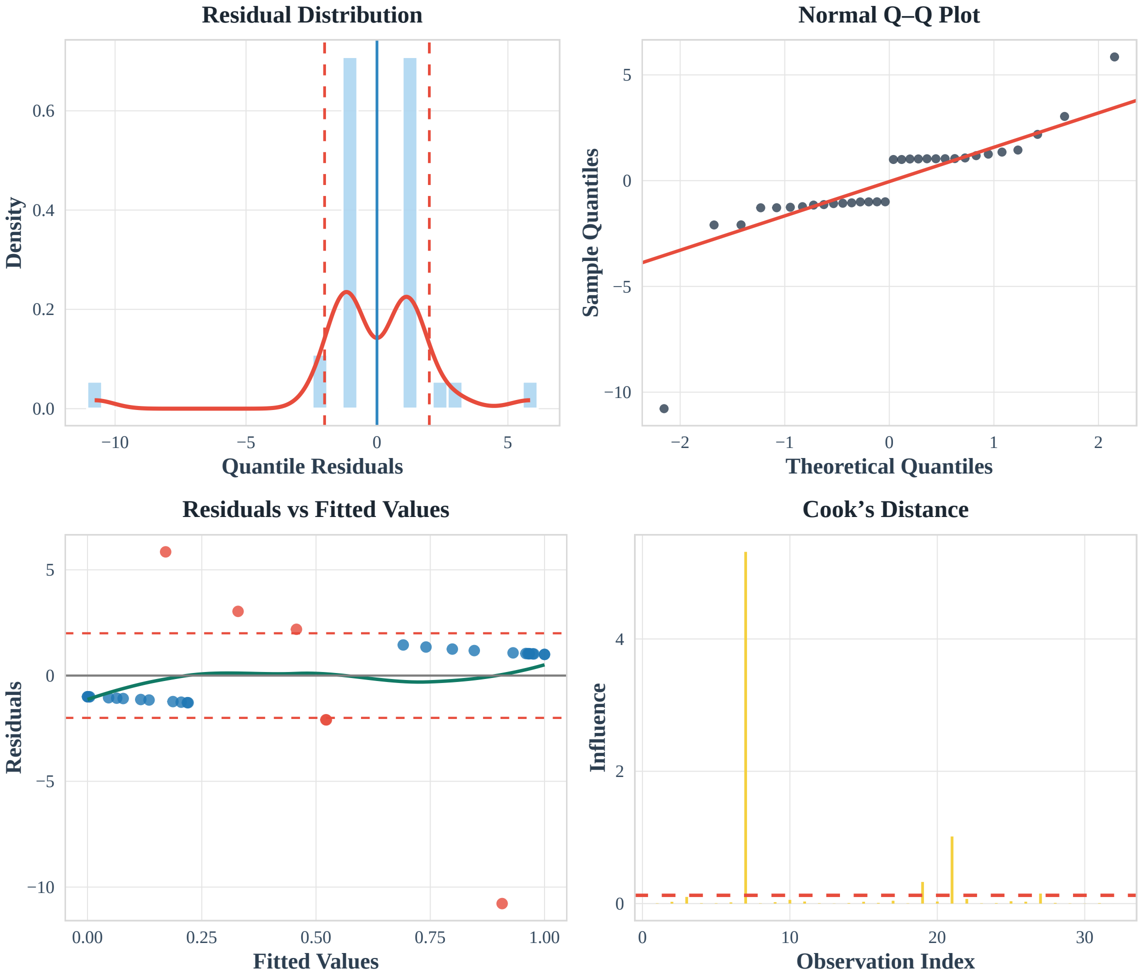


Figure 5. Diagnostic plots for the LoRM in the data set.

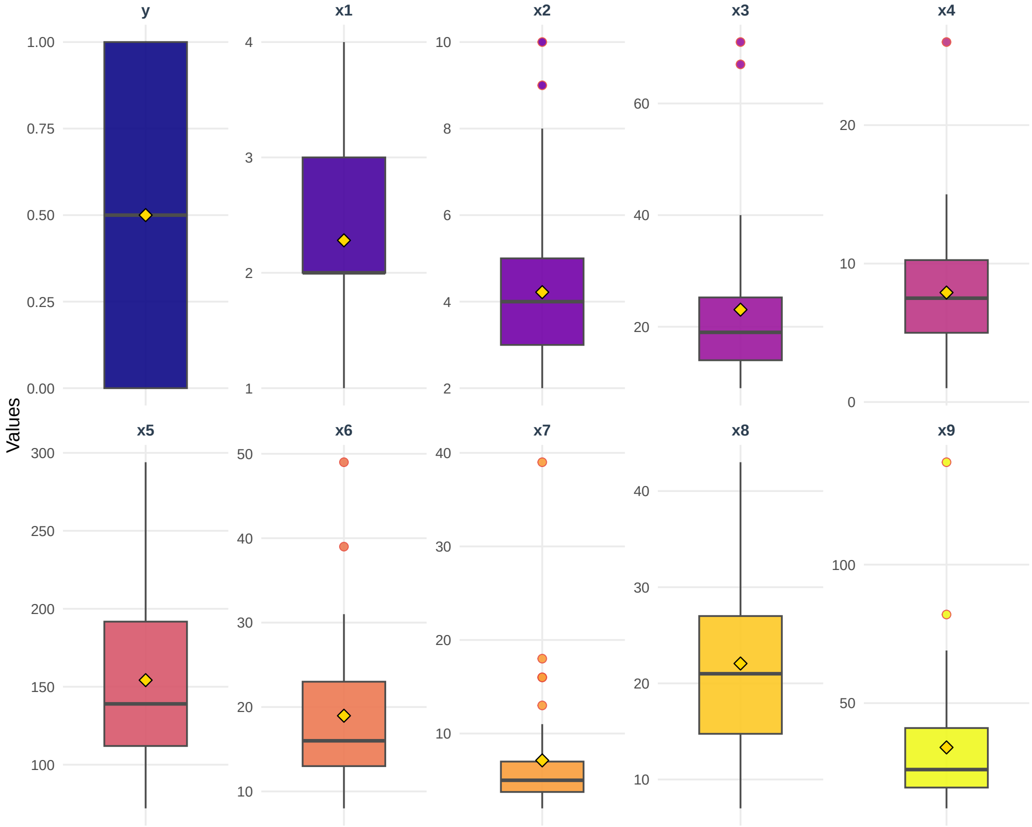


Figure 6. Boxplots of variables in the data set.

Diagnostic evaluation of the data shows that the dataset simultaneously suffers from severe multicollinearity and outliers, which are deleterious to the MLE. First, there exists high multicollinearity among the explanatory variables (metal concentrations). The correlation matrix shows high correlation coefficients among Co (x_2), Cr (x_3), Ni (x_6), and Zn (x_9), and the variance-inflation factors are larger than 10 for Co, Cr, Cu, Ni, and Zn, making it difficult to estimate each regression coefficient under the MLE. Second, these metal concentrations contain outliers. Boxplots suggest extreme values in Co, Cr, Cu, Ni, Pb, and Zn that probably represent real contamination spots rather than measurement error. Outliers with Cook's distance larger than $4/n$ and residuals with absolute values larger than 2 confirm influential observations. Crucially, multicollinearity and outliers exist among the same metal concentrations. The two problems act synergistically to aggravate each other; contamination values are most detrimental when the predictors are highly correlated. This also impacts the binary response variable y (landfill source), as these metals are precisely the ones that differentiate between the two sources. Given the nature of these data, the developed robust ridge-type estimator is appropriate, since the ridge component stabilizes the estimates under multicollinearity and the robust component filters out the negative impact of outliers. Hence, the data becomes a real test of usefulness.

Table 9 illustrates the estimated coefficients (β_0 to β_9) as well as the MSE values corresponding to different estimators applied to the dataset. The estimators include the classical MLE as well as some biased or robust estimators such as the LRRE, LRR-BYE, LRR-CE, LRTE, LRT-BYE, and LRT-CE. These biased estimators are designed to improve the estimation process in the presence of outliers and multicollinearity.

Table 9. MSE values and coefficients of robust and non-robust biased estimators in the data set.

Coefficient	Non-Robust			Robust				Non-Robust		Robust			
	MLE	LRRE		LRR-BYE		LRR-CE		LRTE		LRT-BYE		LRT-CE	
	-	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1$	$\hat{k}_2$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$	$\hat{k}_1^*, \hat{d}_1^*$	$\hat{k}_2^*, \hat{d}_2^*$
β_0	-3.04397	-2.09959	-2.04542	-2.01877	-1.96507	-1.67853	-1.62937	-1.57608	-1.42573	-1.50192	-1.35632	-1.21689	-1.15276
β_1	1.26916	1.00114	0.98417	0.9676	0.95064	0.86858	0.85202	-0.69772	-1.29165	-0.66858	-1.22951	-0.50599	-1.04739
β_2	-1.96053	-1.20034	-1.15813	-1.1478	-1.10608	-1.03177	-0.99257	-1.05336	-1.2193	-1.00328	-1.16079	-0.8242	-0.99257
β_3	0.27985	0.22665	0.22352	0.21945	0.21634	0.12859	0.12567	0.11525	0.0665	0.11194	0.06213	0.04124	0.02315
β_4	0.14114	0.211	0.21494	0.20814	0.21205	0.24578	0.24916	0.06873	-0.54608	0.07007	-0.52219	0.11869	-0.45521
β_5	0.02499	0.01461	0.01403	0.01393	0.01335	0.00988	0.00935	0.01679	0.08284	0.01602	0.08015	0.01102	0.07405
β_6	-8e-05	-0.07951	-0.08383	-0.08003	-0.08428	-0.04776	-0.05174	-0.04885	-0.16183	-0.05082	-0.15869	-0.03477	-0.14553
β_7	0.45933	0.41743	0.41493	0.406	0.40351	0.37347	0.37121	0.2882	-0.35494	0.27969	-0.34193	0.2712	-0.30535
β_8	0.13307	0.12998	0.12997	0.12685	0.12685	0.16129	0.16132	0.22031	-0.14716	0.21286	-0.14387	0.22699	-0.12702
β_9	-0.29445	-0.2717	-0.27032	-0.2644	-0.26303	-0.24905	-0.24769	-0.12888	0.25267	-0.12584	0.24155	-0.12956	0.20915
MSE	12.24952	8.58946	8.51245	8.46463	8.38558	7.97989	7.8962	7.32773	6.94771	7.17476	7.06628	6.61303	6.85871

The MSE values indicate that the regularized robust estimators typically outperform the MLE. For example, the LRR-BYE and LRT-CE have the lowest MSE values. This implies that these estimators provide accurate and stable estimates. These values illustrate the usefulness of our strong shrinkage methods in the presence of outliers and multicollinearity. The different estimators produce different estimates for the coefficients. In the presence of outliers or multicollinearity, the MLE can produce estimates that are too high or too low. In contrast, the robust regularized estimators tend to produce estimates that are more central. This central tendency decreases the variability in the estimates without introducing significant bias, which explains their lower MSE. Table 9 illustrates that implementing robust biased estimators, particularly LRT-CE with ($\hat{k}_1^*$, $\hat{d}_1^*$), results in regression estimates that are more precise and consistent, especially in the presence of multicollinearity and data outliers. These methods have smaller MSE and more reasonable coefficient estimates, which demonstrates that they outperform the classical MLE in practical situations.

Figure 7 shows how estimators perform when the (k) and (d) shrinkage parameters change. It shows a clear bias-variance tradeoff, where all biased estimators have a much lower MSE than the MLE as k increases. The suggested robust biased estimators, especially the LRT-CE, consistently yield the lowest MSE, particularly when d is high (for example, $d \to 1$). The findings indicate that an increase d improves variance reduction and adjusts the optimal k to lower values, suggesting that k and d shrinkage are complementary mechanisms.

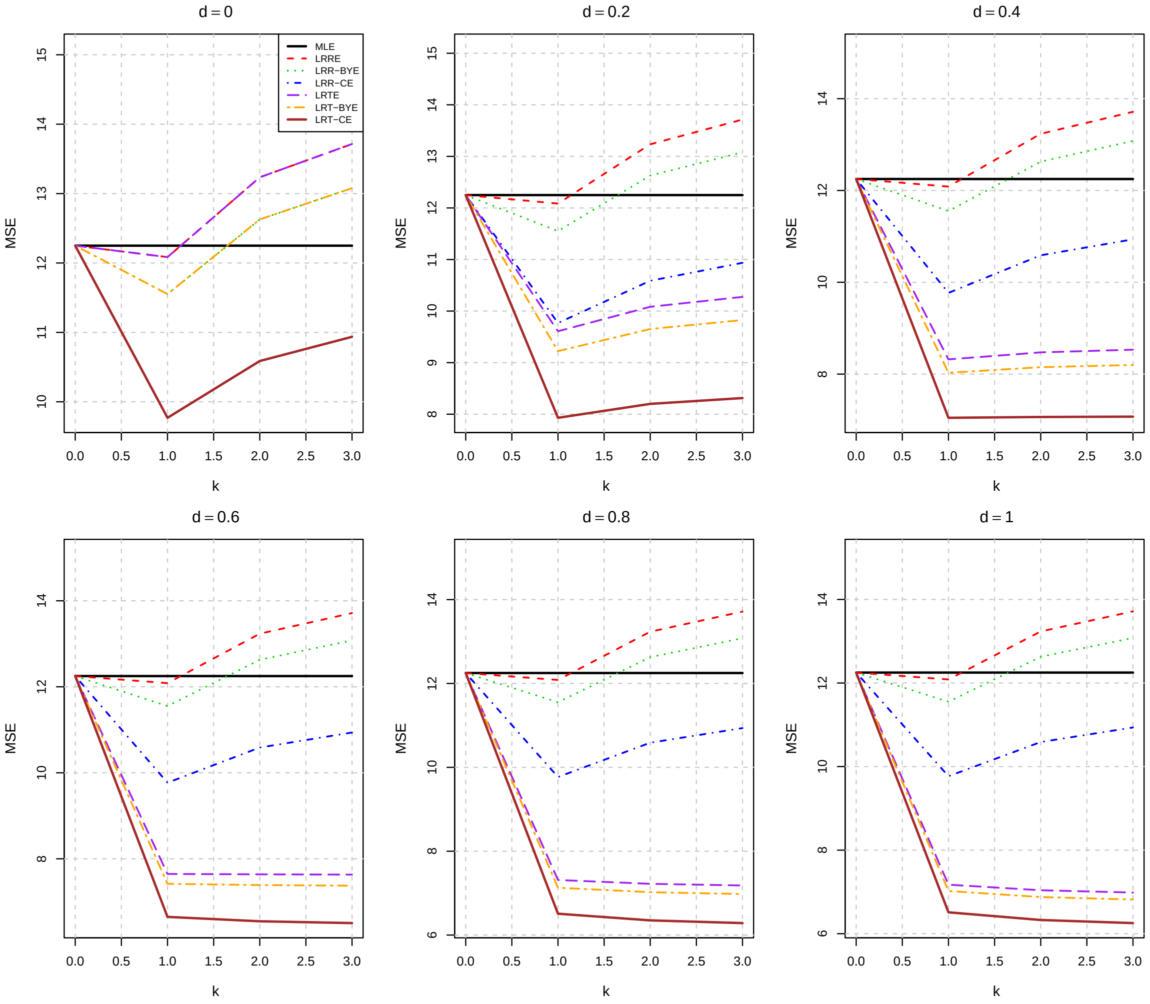


Figure 7. Comparison of MSE among robust and non-robust estimators across different values of the shrinkage parameters (k, d).

6. Conclusions

This study addressed two simultaneous challenges in the LoRM: Multicollinearity among explanatory variables and the presence of influential outliers, which can substantially degrade prediction accuracy. To overcome this limitation, we proposed a class of estimators that integrates robust estimation methods (BYE and CE) with the Liu-type ridge estimator (LRTE), enabling the joint handling of both issues. Monte Carlo simulations and an empirical application demonstrated that the proposed LRT-BYE and LRT-CE estimators consistently outperform existing alternatives. Specifically, the LRT-CE yielded the lowest mean squared error across nearly all simulated scenarios and exhibited superior resistance to the combined effects of multicollinearity and outliers. From a practical perspective, the proposed estimator is particularly beneficial in domains where binary outcome data are frequently accompanied by correlated predictors and potential anomalies, such as in environmental studies, medical research, and financial risk assessment. In such settings, the LRT-CE provides a more reliable alternative to conventional logistic regression. Nonetheless, there are some limitations. The performance of the estimators relies on the selection of shrinkage parameters, which are data-driven. This selection may necessitate the use of more complex, time-consuming cross-validation. The current study presumes that outliers constitute 10% and 20% of the sample, and the method's performance with high-leverage outlier contamination of greater than 30% has not been considered. Each of these limitations has the potential to inform future work. Enhancement of the proposed method is likely with the generalization of the outlined technique to other generalized linear models; the advancement of more goal-adaptive, tuning parameter selection frameworks; and the improvement of the proposed method in high-dimensional cases. Also, other theoretical properties, such as consistency and efficiency of the proposed estimator, are important to discuss.

Author contributions

Eslam Hussam: conceptualization, methodology, formal analysis, software, validation, funding acquisition, writing–original draft, writing–review & editing; Yousef Alharbi: conceptualization, methodology, formal analysis, software, validation, writing–original draft, writing–review & editing; Ahmed M. Gemeay: conceptualization, methodology, formal analysis, software, validation, writing–original draft, writing–review & editing; Samirah Alzubaidi: conceptualization, methodology, formal analysis, software, validation, writing–original draft, writing–review & editing; M. H. Harpy: conceptualization, methodology, formal analysis, software, validation, writing–original draft, writing–review & editing; Ramy Aldallal: conceptualization, methodology, formal analysis, software, validation, writing–original draft, writing–review & editing; M. S. Mohamed: conceptualization, methodology, formal analysis, software, validation, writing–original draft, writing–review & editing; Ali T. Hammad: conceptualization, methodology, formal analysis, software, validation, writing–original draft, writing–review & editing. All authors have read and agreed to the published version of the manuscript.

Use of Generative-AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2025/01/36686).

Conflicts of interest

The authors declare no conflicts of interest.

References

1. T. G. Nick, K. M. Campbell, Logistic regression, In: W. T. Ambrosius, *Topics in biostatistics*, Springer, 2007, 273–301. https://doi.org/10.1007/978-1-59745-530-5_14

2. M. Jain, A. Srihari, Comparison of machine learning algorithm in intrusion detection systems: a review using binary logistic regression, *Authorea Preprints*, 2025.
3. H. Hocek, S. Yay, D. Yazir, Comprehensive analysis of ship detention probabilities using binary logistic regression method with machine learning, *Ocean Eng.*, **315** (2025), 119889. https://doi.org/10.1016/j.oceaneng.2024.119889
4. D. Dey, M. S. Haque, M. M. Islam, U. I. Aishi, S. S. Shammy, M. S. A. Mayen, et al., The proper application of logistic regression model in complex survey data: a systematic review, *BMC Med. Res. Methodol.*, **25** (2025), 15. https://doi.org/10.1186/s12874-024-02454-5
5. H. Hasim, M. Salam, A. A. Sulaiman, M. H. Jamil, H. Iswoyo, P. Diansari, et al., Employing binary logistic regression in modeling the effectiveness of agricultural extension in clove farming: facts and findings from sidrap regency, Indonesia, *Sustainability*, **17** (2025), 2786.
6. M. N. Akram, M. Amin, A. Elhassanein, M. A. Ullah, A new modified ridge-type estimator for the beta regression model: simulation and application, *AIMS Math.*, **7** (2022), 1035–1057. https://doi.org/10.3934/math.2022062
7. R. Alharbi, A. S. Alghamdi, A new almost unbiased estimator for beta regression model under multicollinearity, *AIMS Math.*, **11** (2026), 85–126. https://doi.org/10.3934/math.2026005
8. R. L. Schaefer, L. D. Roi, R. A. Wolfe, A ridge logistic estimator, *Commun. Stat.*, **13** (1984), 99–113. https://doi.org/10.1080/03610928408828664
9. K. Månsson, B. M. G. Kibria, G. Shukur, On Liu estimators for the logit regression model, *Econ. Model.*, **29** (2012), 1483–1488. https://doi.org/10.1016/j.econmod.2011.11.015
10. M. K. Çetinkaya, A new estimator for the multicollinear logistic regression model, *Stat. Pap.*, **66** (2025), 55. https://doi.org/10.1007/s00362-025-01675-0
11. O. J. Oladapo, O. O. Alabi, K. Ayinde, Another new two parameter estimator in dealing with multicollinearity in the logistic regression model, *Int. J. Math. Sci. Optim.*, **10** (2024), 22–35. https://doi.org/10.5281/zenodo.10937145
12. A. F. Lukman, B. M. G. Kibria, C. K. Nziku, M. Amin, E. T. Adewuyi, R. Farghali, KL estimator: dealing with multicollinearity in the logistic regression model, *Mathematics*, **11** (2023), 340. https://doi.org/10.3390/math11020340
13. F. A. Awwad, K. A. Odeniyi, I. Dawoud, Z. Y. Algamal, M. R. Abonazel, B. M. G. Kibria, et al., New two-parameter estimators for the logistic regression model with multicollinearity, *WSEAS Trans. Math.*, **21** (2022), 403–414. https://doi.org/10.37394/23206.2022.21.48
14. P. Rousseeuw, V. Yohai, Robust regression by means of S-estimators, In: J. Franke, W. Härdle, D. Martin, *Robust and nonlinear time series analysis*, Springer, 1984, 256–272. https://doi.org/10.1007/978-1-4615-7821-5_15
15. J. Feng, H. Xu, S. Mannor, S. Yan, Robust logistic regression and classification, *Adv. Neural Inf. Process. Syst.*, **27** (2014), 253–261.
16. A. M. Bianco, V. J. Yohai, Robust estimation in the logistic regression model, In: H. Rieder, *Robust statistics, data analysis, and computer intensive methods*, Springer, 1996, 17–34. https://doi.org/10.1007/978-1-4612-2380-1_2

17. R. J. Carroll, S. Pederson, On robustness in the logistic regression model, *J. R. Stat. Soc.*, **55** (1993), 693–706. https://doi.org/10.1111/j.2517-6161.1993.tb01934.x
18. A. Hakimi, A. Amiri, R. Kamranrad, Robust approaches for monitoring logistic regression profiles under outliers, *Int. J. Qual. Reliab. Manage.*, **34** (2017), 494–507. https://doi.org/10.1108/IJQRM-04-2015-0053
19. B. Akturk, U. Beyaztas, H. L. Shang, A. Mandal, Robust functional logistic regression, *Adv. Data Anal. Classif.*, **19** (2025), 121–145. https://doi.org/10.1007/s11634-023-00577-z
20. A. M. Bianco, G. Boente, G. Chebi, Penalized robust estimators in sparse logistic regression, *Test*, **31** (2022), 563–594. https://doi.org/10.1007/s11749-021-00792-w
21. B. Shin, S. Lee, Robust logistic regression with shift parameter estimation, *J. Stat. Comput. Simul.*, **93** (2023), 2625–2641. https://doi.org/10.1080/00949655.2023.2201008
22. I. A. I. Ahmed, W. Cheng, The performance of robust methods in logistic regression model, *Open J. Stat.*, **10** (2020), 127. https://doi.org/10.4236/ojs.2020.101010
23. A. Ostovar, D. D. Davari, M. Dzikuc, Determinants of design with multilayer perceptron neural networks: a comparison with logistic regression, *Sustainability*, **17** (2025), 2611. https://doi.org/10.3390/su17062611
24. A. F. Lukman, S. Mohammed, O. Olaluwoye, R. A. Farghali, Handling multicollinearity and outliers in logistic regression using the robust Kibria–Lukman estimator, *Axioms*, **14** (2024), 19. https://doi.org/10.3390/axioms14010019
25. K. C. Arum, F. I. Ugwuowo, H. E. Oranye, T. O. Alakija, T. E. Ugah, O. C. Asogwa, Combating outliers and multicollinearity in linear regression model using robust Kibria-Lukman mixed with principal component estimator, simulation and computation, *Sci. Afr.*, **19** (2023), e01566. https://doi.org/10.1016/j.sciaf.2023.e01566
26. H. H. Mohammad, A. T. Hammad, A. A. El-Helbawy, Z. I. Kalantan, A. Habineza, E. Hussam, et al., New robust two-parameter estimator for overcoming outliers and multicollinearity in Poisson regression model, *Sci. Rep.*, **15** (2025), 27445. https://doi.org/10.1038/s41598-025-12646-8
27. A. T. Hammad, I. Elbatal, E. M. Almetwally, M. M. A. El-Raouf, M. A. El-Qurashi, A. M. Gemeay, A novel robust estimator for addressing multicollinearity and outliers in Beta regression: simulation and application, *AIMS Math.*, **10** (2025), 21549–21580. https://doi.org/10.3934/math.2025958
28. A. M. Alshangiti, A. M. Gemeay, M. E. Bakr, A. M. Yousuf, M. A. El-Qurashi, O. S. Balogun, et al., New robust estimator for handling outliers and multicollinearity in gamma regression model with application to breast cancer data, *Sci. Rep.*, **15** (2025), 38436. https://doi.org/10.1038/s41598-025-25231-w
29. W. B. Altukhaes, M. Roozbeh, N. A. Mohamed, Feasible robust Liu estimator to combat outliers and multicollinearity effects in restricted semiparametric regression model, *AIMS Math.*, **9** (2024), 31581–31606. https://doi.org/10.3934/math.20241519
30. D. Pregibon, Resistant fits for some commonly used logistic models with medical applications, *Biometrics*, **38** (1982), 485–498. https://doi.org/10.2307/2530463

31. H. R. Künsch, L. A. Stefanski, R. J. Carroll, Conditionally unbiased bounded-influence estimation in general regression models, with applications to generalized linear models, *J. Amer. Stat. Assoc.*, **84** (1989), 460–466. https://doi.org/10.1080/01621459.1989.10478791
32. A. E. Hoerl, R. W. Kennard, Ridge regression: biased estimation for nonorthogonal problems, *Technometrics*, **12** (1970), 55–67.
33. A. F. Lukman, K. Ayinde, S. Binuomote, O. A. Clement, Modified ridge-type estimator to combat multicollinearity: Application to chemical data, *J. Chemom.*, **33** (2019), e3125. https://doi.org/10.1002/cem.3125
34. B. M. G. Kibria, A. F. Lukman, A new ridge-type estimator for the linear regression model: simulations and applications, *Scientifica*, **2020** (2020), 9758378. https://doi.org/10.1155/2020/9758378
35. G. Khalaf, G. Shukur, Choosing ridge parameter for regression problems, *Commun. Stat.*, **34** (2005), 1177–1182. https://doi.org/10.1081/STA-200056836
36. A. T. Hammad, E. H. Hafez, U. Shahzad, E. Yıldırım, E. M. Almetwally, B. M. G. Kibria, New modified Liu estimators to handle the multicollinearity in the beta regression model: simulation and applications, *Mod. J. Stat.*, **1** (2025), 58–79. https://doi.org/10.64389/mjs.2025.01111
37. O. T. Olaluwoye, A. F. Lukman, M. A. Alrasheedi, W. N. Nzomo, R. A. Farghali, Robust estimation methods for addressing multicollinearity and outliers in beta regression models, *Sci. Rep.*, **15** (2025), 11649. https://doi.org/10.1038/s41598-025-85553-7
38. A. T. Hammad, I. Elbatal, O. A. Alqasem, E. M. Almetwally, A. M. A. Haleeb, A. M. Yousuf, et al., New robust ridge-type estimators for beta regression under multicollinearity and outliers with application to breast cancer data, *Appl. Math. Sci. Eng.*, **33** (2025), 2590432. https://doi.org/10.1080/27690911.2025.2590432
39. T. Alharbi, A. S. El-Sorogy, N. Rikan, A GIS and multivariate analysis approach for mapping heavy metals and metalloids contamination in landfills: a case study from Al-Kharj, Saudi Arabia, *Land*, **14** (2025), 1697. https://doi.org/10.3390/land14081697

AIMS Press