Robust sparse principal component analysis: situation of full sparseness

B. BARIŞ ALKAN AND I. ÜNALDI

Abstract

Principal Component Analysis (PCA) is the main method of dimension reduction and data processing when the dataset is of high dimension. Therefore, PCA is a widely used method in almost all scientific fields. Because PCA is a linear combination of the original variables, the interpretation process of the analysis results is often encountered with some difficulties. The approaches proposed for solving these problems are called to as Sparse Principal Component Analysis (SPCA). Sparse approaches are not robust in existence of outliers in the data set. In this study, the performance of the approach proposed by Croux et al. (2013), which combines the advantageous properties of SPCA and Robust Principal Component Analysis (RPCA), will be examined through one real and three artificial datasets in the situation of full sparseness. In the light of the findings, it is recommended to use robust sparse PCA based on projection pursuit in analyzing the data. Another important finding obtained from the study is that the BIC and TPO criteria used in determining lambda are not much superior to each other. We suggest choosing one of these two criteria that give an optimal result.

Mathematics Subject Classification 2000: 62hxx, 62-07
Key Words: Principal Component, Outliers, Sparsity, Robustness.

1. INTRODUCTION

Principal component analysis (PCA) is the first referenced dimension reduction method when there is a multidimensional data set. PCA aims to reduce size by finding k $(k \leq p)$ linear combinations of the p original variables in a reduced-dimensional space. Thus, more easily interpretable results are obtained. Principal components (PCs) are vectors that make the maximum variance of k linear combinations (Croux et al., 2013). The first k principal component explains an important part of the total variance of the original data. The PCA uses the sample covariance (or correlation) matrix and the mean vector. Although PCA is a widely used multivariate statistical analysis method in almost all scientific fields, it has too many deficiencies.

In PCA, there is a transformation matrix that identifies the principal components It is called the *loading matrix*. Interpretation of loadings is often a difficult process. Various suggestions have been put forward in the literature to solve the problem of interpretability of PCA results. One of these suggestions is to use the rotation methods

10.2478/jamsi-2022-0001

to obtain the simple structure of the loading matrix (Jolliffe, 1995). Researchers, rather than the contribution of a variable with moderate loading values, which contribute a weak or uncertain contribution to a principal component, want to be concerned with the contribution of variables that have the big loadings. Otherwise, interpreting principal components can be very difficult. In order to increase the interpretability of PCA results, many approaches, called sparse principal component analysis (SPCA), are presented to estimate the principal components with many zero loadings. When the dataset is multidimensional, the SPCA is useful because only a subset of variables needs to be analyzed. SPCA offers researchers the advantage of easier interpretation of principal components. The easiest way to do this is to set the loadings to zero that have smaller values than a certain threshold. This method is called simple thresholding. Cadima and Jolliffe (1995) found that this method could be misleading. In this context, they emphasized the necessity of examining the standard deviations of variables for determining the contribution of a variable to a given principal component. In the context of PCA, the concept of sparseness was first described by Jolliffe et al. (2003). Later, Zou et al. (2006) suggested SPCA algorithm based on elastic-net regression, which gives better results than SCoTLASS. The adaptations and improvements of the method were made by Wang et al. (2009). D'Aspremont et al. (2007) suggested a direct formulation for SPCA. Sigg et al. (2008) designed EM-PCA based on probabilistic PCA to solve SPCA. Witten et al. (2009) developed a general procedure for the separation of penalty matrix. They showed how this procedure is applied. Journée et al. (2010) derived the GPower algorithm that formulated the SPCA. Guo et al. (2010) have introduced a fusion penalty that captures block structures within the variables. Other recommendations for obtaining SPCA are described in Jenatton et al. (2010) and Bien et al. (2010).

In the presence of outliers in the dataset, many robust alternative solutions have been proposed for PCA. The most important of these were Li and Chen (1985), Hubert et al. (2002), projection pursuit PCA approach given by Croux and Ruiz-Gazen (2005), global PCA approach given by Locantore et al. (1999) as a robust PCA approach. This approach works well as long as there is a robust estimate of multivariate location and scale is possible.

The robust PCA approach proposed by Croux and Ruiz-Gazen (2005) based on the projection pursuit, is suitable when the number of variables is more than the number of observations and the presence of high dimensional data sets. The advantage of the projection pursuit method is that it is not necessary to estimate the covariance matrix for PCA. Robust estimates of eigenvalues and eigenvectors are obtained successively. In the literature, it is possible to find many studies about the robust versions of the PCA and their robustly characteristics. In addition, sparse PCA has also gained importance in the analysis of multidimensional (large data) datasets, which have emerged in recent years as a result of increasingly rapid technological developments, especially in areas such as internet technologies, gene expression and machine learning. He et al. (2014) combining the advantageous properties of Croux et al. (2013) sparse and robust PCA, proposed an algorithm for improving the durability of SPCA. They examined the robustness and effectiveness of the proposed algorithm through artificial and real data sets. A recent study on SPCA was made by Hubert et al. (2016). The main difference from Hubert et al. (2016) 's study of the Croux et al. (2013) is the separation of the sparse step and the steps of identifying outlier observations.

In this study, the performance of the approach proposed by Croux et al. (2013), which combines the important properties of SPCA and Robust Principal Component Analysis (RPCA), will be examined through one real and three artificial datasets in the situation of full sparseness. In the light of the findings, in cases where the number of variables (p) is greater than the number of observations (n), it is recommended to use robust sparse PCA based on projection pursuit in analysing the data. Another important conclusion obtained from the study is that the BIC and TPO criteria used in determining λ are not much superior to each other. We propose to choose the criterion that gives an optimal result. In the second part of the study, the mathematical theory required for SPCA and robust sparse PCA is summarized. In the next section, the applications of the methods discussed in detail in the previous sections on real and artificial datasets are given. In the last section, the results obtained from the study are discussed.

2. SPARSE PCA (SPCA)

The first PCA vector for the $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n \in \mathbb{R}^p$, n multivariable observations in the rows of the X data matrix is given by,

$$\operatorname{argmax}_{\|\boldsymbol{p}\|=1} V(\boldsymbol{p}'x_1, \dots, \boldsymbol{p}'x_n) \tag{1}$$

Here, V is a measure of variance. In the standard non-robust state V is the experimental variant (var) and the optimal $\boldsymbol{p}_1$ vector corresponds to the first eigenvector of the sample covariance matrix. Equation (1) is the projection pursuit formulation to find the first PC.

Robust PCA vectors can be easily obtained by taking a robust variance measure for V. This measure can be selected as a quadratic median absolute deviation or as a more efficient quadratic $Q_n$estimator.

The Q_n scale estimator is defined as the first quarter of all binary distances, $|y_i - y_j|$, $1 \leq i < j \leq n$, for a univariate data set $y_1,\dots,y_n$. This first quarter should then be multiplied by the constant 2.219 to obtain a consistent estimator for a normal distribution scale. Therefore, Q_n^2 is a consistent and robust predictor of variance. Croux and Ruiz-Gazen (2005) suggested using the Q_n^2 estimator as the projection tracking index which provides robust and effective estimates for the basic components. In this study, Q_n^2 was taken as a robust variance estimator.

Suppose the first $j-1$ PCA vectors are present (j >1). In this case j. vector ($j \leq p$) is defined as,

$$\operatorname{argmax}_{\|\boldsymbol{p}\|=1, \boldsymbol{p}\perp\boldsymbol{p}_1,\dots,\boldsymbol{p}\perp\boldsymbol{p}_{j-1}} V(\boldsymbol{p}'x_1, \dots, \boldsymbol{p}'x_n). \tag{2}$$

The sparsity of PCA was first used with SCoTLASS method by Jolliffe et al. (2003). SCoTLASS combines an L_1 constraint with PCA to provide sparse loadings. The resulting objective function is written as

$$max_{\|\boldsymbol{p}\|=1, \boldsymbol{p}\perp\boldsymbol{p}_1,\dots,\boldsymbol{p}\perp\boldsymbol{p}_{j-1}} \boldsymbol{p}'\boldsymbol{S}\boldsymbol{p}, \quad \text{constraint } \|\boldsymbol{p}\|_1 \leq \boldsymbol{t}, \tag{3}$$

This objective function tries to find the vertical loadings $\boldsymbol{p}_j$that maximize the variance described. Here, $\|\boldsymbol{p}_j\|_1$ is the L_1 norm of $\boldsymbol{p}_j$, $\|\boldsymbol{p}\|_1 = \sum_{j=1}^{p} |\boldsymbol{p}_j|$. The dual formulation of the above form is given by

$$max_{\|\boldsymbol{p}\|=1,\boldsymbol{p}\perp\boldsymbol{p}_1,\dots,\boldsymbol{p}\perp\boldsymbol{p}_{j-1}}(\boldsymbol{p}'\boldsymbol{S}\boldsymbol{p}-\lambda_1\|\boldsymbol{p}\|_1) \quad (4)$$

Where $\boldsymbol{p_j}$ is the $\boldsymbol{j}$th PCA vector. $\boldsymbol{\lambda_1}$ is the sparsity parameter for SCoTLASS and is used instead of $\boldsymbol{t}$. The larger values of $\boldsymbol{\lambda_j}$ mean greater sparsity and the zero value means no sparsity.

2.1. Robust Sparse PCA

Croux et al. (2013) proposed the robust SPCA method, which combines the projection pursuit (PP) approach and SPCA. Their approach consists of adding L_1 penalty to PP equations. This method searches for vectors that maximize the scale of the data projected onto these vectors under the constraint that vector loadings are not too large. The principal components are obtained directly using the PP approach without estimating a covariance matrix.

To find the first sparse PCA vector, the L_1 constraint is added and

$$\tilde{\boldsymbol{p}}_1=\text{argmax}_{\|\boldsymbol{p}\|=1}(V(\boldsymbol{p}'x_1,\dots,\boldsymbol{p}'x_n)-\lambda_1\|\boldsymbol{p}\|_1) \quad (5)$$

equation is obtained. Here, the vector $\tilde{\boldsymbol{p}}_1$ is the first SPCA vector, its sparsity is controlled by parameter λ_1. If λ_1= 0, the constraint is not added to the first PCA vector $\boldsymbol{p}_1$. Similarly, jth SPCA vector ($1 < j \leq p$),

$$\tilde{\boldsymbol{p}}_j=\text{argmax}_{\|\boldsymbol{p}\|=1,\boldsymbol{p}\perp\tilde{\boldsymbol{p}}_1,\dots,\boldsymbol{p}\perp\tilde{\boldsymbol{p}}_{j-1}}(V(\boldsymbol{p}'x_1,\dots,\boldsymbol{p}'x_n)-\lambda_j\|\boldsymbol{p}\|_1) \quad (6)$$

as defined. Here, V is the variance estimator, a measure of the scale. This is the empirical (experimental) variant for the classical basic components. For the robust sparse PCA, this variance estimator is a robust variance estimator such as squared Q_n. If $V=var$ exist, then equality (4) and equality (5) are the same. Croux et al. (2013), Rousseeuw and Croux (1993) proposed the robust Q_n estimator. This Q_n estimator is the first quartile of the binary distances between the elements of a vector.

It is not easy to solve the optimization problems given in equations (5) and (6). Croux et al. (2013) is an algorithm based on iterative grid searches in spaces stretched by binary pairs of variables to find sparse vectors. Croux et al. (2007), the grid algorithm used to obtain PCA vectors is a well-working algorithm. This algorithm was developed by Croux et al. (2013) expanded to sparse PCA.

Grid algorithm is a kind of coordinate origin methods. The extended grid algorithm always provides robust and sparse solutions with an appropriate calculation time (Croux et al., 2013).

2.2. Selection of Sparsity Parameter λ

The selection of the sparsity parameter λ is made by optimizing the objective function calculated for a set of n. λ can take different values, ranging from zero to the maximum value. If the maximum value that can be taken by λ cannot be determined, then the minimum λ value that provides full sparseness is preferred. The absolute sparseness state refers to the minimum absolute sum of the loadings, which usually consist only of zero and one in the loadings matrix. Researchers can complete this process using one of two optimization approaches for λ selection. These approaches are given as TPO (Tradeoff Product Optimization) and BIC (Bayes Information Criteria) respectively. The BIC approach selects the same λ value for all PCs and expresses a specific choice of the number of PCs considered. However, the TPO approach is optimized for each PC separately, so that different λ values are derived in a model that is not bound to a decision over k. In the TPO approach, tradeoff in the context of SPCA refers to the sparsity gained against the explained variance loss. The TPO maximizes the product of the variance explained by the number of zero loadings of the sparse principal component jth. Here, λ_j values for k PC according to TPO approach are obtained by equation (7).

$$\lambda_j{}^{TPO} = V\big((STB)_j\big)\alpha_j,\ \ j = 1,2,\ldots,k \tag{7}$$

where α_j is the number of zero loadings for the jth SPC. Croux et al. (2013) used a BIC type criterion to select the λ. Let P_k^s and P_k^c be the matrices of loadings that contain sparse and non-sparse basic components that contain the first k PC, respectively. From this, residual matrices $R^s = X - XP_k^s(P_k^s)'$ and $R^c = X -$

$XP_k^c(P_k^c)'$are defined. The jth columns of residual matrices are defined as $r_j^s = (r_{1j}^s, r_{2j}^s, \ldots, r_{nj}^s)$ and $r_j^c = (r_{1j}^c, r_{2j}^c, \ldots, r_{nj}^c)$ respectively.

The BIC criterion is defined by the equation

$$\lambda^{BIC} = \frac{\sum_{j=1}^{k} V(r_j^s)}{\sum_{j=1}^{k} V(r_j^c)} + \beta \frac{\log n(n)}{n} \tag{8}$$

Here, β is the number of non-zero loadings.

In practice, the selection of λ is made by minimizing λ in a grid $[0, \lambda_{max}]$ range. Here, λ_{max} is the result of the full sparseness of k-component sparse PCA results. In addition to λ, it is also necessary to select the number of components of k. The appropriate choice of k is an old and common problem in PCA. Many suggestions have been made to solve this problem. In this study, the number of k is determined by the robust variance explanation ratio (VER),

$$VER_k = \frac{V(T_k)}{V(X)} \tag{9}$$

Here, T_k contains the principal component scores. For $V = var$, VER_k, k is the ratio of the sum of the biggest eigenvalues to the sum of all eigenvalues of the sample covariance matrix. For this value of k, a selected λ must have a more sparse matrix and a smaller robust variance.

3. ARTIFICIAL AND REAL DATA EXAMPLES

In this section, the performance of RPCA based on PP (PP-RPCA) nd robust sparse PCA based on PP (PP-RSPCA) methods based on projection-pursuit (PP) will be compared over one real and three artificial data sets. The codes written with the help of pcaPP, pls, lars, lasso2, spls, lattice, elasticnet, stats, stats4, robust, robustbase, rrcov, rrcovHD libraries were used in the R program to apply PP-RPCA and PP-RSPCA to real and artificial data sets (R Core Team,2019).

Artificial data sets were used in the design given in the study. As the actual data set was made for the monthly stock exchange traded on the Istanbul Stock Exchange between January 2005 and March 2013.

Artificial Data Designs

1) Artificial Data Set 1

10% outlier situation: A data set of 100 observations (n = 100) and 5 variables (p = 5) was generated from a multivariate t distribution (T_5):10% and $N_5(0, I)$:90%.

20% outlier situation: A data set of 100 observations (n = 100) and 5 variables (p = 5) was generated from a multivariate t distribution (T_5): 20% and $N_5(0, I)$:80%.

30% outlier situation: A data set of 100 observations (n = 100) and 5 variables (p = 5) was generated from a multivariate t distribution (T_5): 30% and $N_5(0, I)$:70%

2) Artificial Data Set 2

It was planned to produce a 215 × 6 data set consisting of 215 observations (n = 215) and 6 variables (p = 6). Of these 215 observations, N_6 ($\mu 1, \Sigma 1$): 200 and N_6 ($\mu 2, \Sigma 2$): 15.

200 observations were generated from the multivariate normal distribution and 15 from the multivariate normal distribution.

$\mu 1 = [0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0]$, $\Sigma 1 = diag\,[5 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \;\; 1]$

$\mu 2 = [0 \;\; 20 \;\; 20 \;\; 20 \;\; 20 \;\; 20]$, $\Sigma 2 = diag\,[1 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \;\; 1]$

3) Artificial Data Set 3 (p > n status)

An 18 × 20 data set consisting of 18 observations and 20 variables was produced. Of the 18 observations, N_{20} ($\mu 3, \Sigma 3$): 15 and N_{20} ($\mu 4, \Sigma 4$): 3.

$\mu 3 = [0 \;\; 0 \;\; 0 \;\; \ldots \;\; 0]$, $\Sigma 3 = diag[5 \;\; 1 \;\; 1 \;\; \ldots \;\; 1]$

$\mu 4 = [0 \;\; 20 \;\; 20 \;\; \ldots \;\; 20]$, $\Sigma 4 = diag[1 \;\; 1 \;\; 1 \;\; \ldots \;\; 1]$

Real Data (Stock Exchange Data)

The data set containing the quarterly return rates of eight stocks traded between January 2005 and March 2013 on the Istanbul Stock Exchange was used. The analyzed dataset has a size of 33 × 8.

3.1. Comparison of PP-RPCA and PP-RSPCA over Artificial Data Set 1

The Artificial Data Set 1 has been analyzed with the written R code and the results are presented in Table 3.1, Table 3.2, Table 3.3 according to 10%, 20%, 30% outlier rates. For the 10% outlier rate, λ's were determined by using TPO and BIC criteria, and the most optimal λ value was selected as $\lambda = 7.8$ according to TPO criteria. When Table 3.1 is examined, it is seen that PP-RPCA explains 88% of the total variance with the first four principal components. When PP-RSPCA results are analyzed, it is seen that 84% of the total variance is explained by the first four main components. In this case, the cost of using the sparsity feature is equivalent to the loss of 4% variance explanation rate. However, in Table 3.1, it is seen that the number of non-zero loadings in the PP-RPCA is 20 while the number of non-zero loadings in the PP-RSPCA is reduced to 4. It is seen that the sparsity feature added to the loadings matrix facilitates the interpretation of the main components with extremely good performance and a low error rate.

Table 3.1. In the presence of 10% outlier of the variables, the first four non sparse robust ($\lambda = 0$) and sparse robust ($\lambda = 7.8$) main component loadings

Variables	PP-RPCA				PP-RSPCA			
	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4
X1	0.59	0.30	-0.59	-0.07	1.00	0.00	0.00	0.00
X2	0.18	-0.09	0.65	0.24	0.00	0.00	0.00	1.00
X3	0.74	0.03	0.30	0.22	0.00	0.00	0.00	0.00
X4	-0.14	0.92	0.31	-0.21	0.00	0.00	1.00	0.00
X5	-0.21	0.25	-0.22	0.92	0.00	1.00	0.00	0.00
VER* %	0.25	0.24	0.23	0.16	0.25	0.21	0.19	0.18
Cumulative VER %	0.25	0.48	0.72	0.88	0.25	0.46	0.66	0.84

For the 20% outlier rate, λ's were determined by using TPO and BIC criteria and $\lambda = 12.14$ was selected as the most optimal λ value according to TPO criteria.

When Table 3.2 is examined, it is seen that PP-RPCA explains 88% of the total variance with the first four main components. When PP-RSPCA results are analyzed, it is seen that 86% of the total variance is explained by the first four main components.

In this case, the cost of using the sparsity feature is equivalent to the loss of 2% variance explanation rate. However, in Table 3.2, it is seen that the number of non-zero loads in PP-RPCA is 19 while the number of non-zero loads in PP-RSPCA is reduced to 4. It is seen that the sparsity feature added to the loadings matrix facilitates the interpretation of the principal components with extremely good performance and a low error rate.

Table 3.2. In the presence of 20% outlier of the variables, the first four non sparse robust ($\lambda = 0$) and sparse robust ($\lambda = 12.14$) main component loadings

Variables	PP-RPCA				PP-RSPCA			
	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4
X1	0.04	0.73	-0.67	0.01	0.00	1.00	0.00	0.00
X2	-0.03	0.57	0.71	0.00	0.00	0.00	1.00	0.00
X3	0.68	-0.27	-0.12	-0.29	0.00	0.00	0.00	1.00
X4	-0.28	-0.24	-0.14	0.74	1.00	0.00	0.00	0.00
X5	0.67	0.15	0.12	0.61	0.00	0.00	0.00	0.00
VER* %	0.31	0.23	0.19	0.15	0.23	0.23	0.21	0.19
Cumulative VER* %	0.31	0.54	0.73	0.88	0.23	0.46	0.67	0.86

For the case of 30% outlier, λ's were determined by using TPO and BIC criteria and $\lambda = 3.87$ according to TPO criterion.

When Table 3.3 is examined, it is seen that PP-SPCA explains 91% of the total variance with the first four main components. When PP-RSPCA results are analyzed, it is seen that 81% of the total variance is explained by the first four main components. In this case, the cost of using the sparsity feature is equivalent to the loss of variance explanation rate of 10%. However, in Table 3.3, it is seen that the number of non-zero loads in PP-RPCA is 19 while the number of non-zero loadings in PP-RSPCA is reduced to 7. It is seen that the sparsity feature added to the loadings matrix facilitates the interpretation of the principal components with extremely good performance and a low error rate.

Table 3.3. In the presence of 30% outlier of the variables, the first four non sparse robust ($\lambda = 0$) and sparse robust ($\lambda = 3.87$) main component loadings

Variables	PP-RPCA				PP-RSPCA			
	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4
X1	0.04	0.73	-0.67	0.01	0.00	1.00	0.00	0.00
X2	-0.03	0.57	0.71	0.00	0.00	0.00	1.00	0.00
X3	0.68	-0.27	-0.12	-0.29	0.00	0.00	0.00	1.00
X4	-0.28	-0.24	-0.14	0.74	1.00	0.00	0.00	0.00
X5	0.67	0.15	0.12	0.61	0.00	0.00	0.00	0.00
VER %	0.31	0.23	0.19	0.15	0.23	0.23	0.21	0.19
Cumulative VER %	0.31	0.54	0.73	0.88	0.23	0.46	0.67	0.86

3.2. Comparison of PP-RPCA and PP-RSPCA over Artificial Data Set 2

The Artificial Data Set 2 was analyzed by R code and the results are presented in Table 3.4.

Table 3.4. The first four non sparse robust ($\lambda = 0$) and sparse robust ($\lambda = 11.03$) principal component loadings of variables in Artificial Dataset 2

Variables	PP-RPCA				PP-RSPCA			
	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4
X1	0.99	-0.06	0.03	0.11	1.00	0.00	0.00	0.00
X2	0.04	0.26	0.86	-0.18	0.00	0.00	0.00	1.00
X3	0.08	0.79	-0.23	-0.47	0.00	0.00	0.00	0.00
X4	0.07	0.00	0.08	-0.18	0.00	0.00	1.00	0.00
X5	0.07	0.40	-0.36	0.39	0.00	0.00	0.00	0.00
X6	-0.11	0.37	0.27	0.74	0.00	1.00	0.00	0.00
VER %	0.23	0.18	0.17	0.16	0.18	0.18	0.18	0.16
Cumulative VER %	0.23	0.41	0.59	0.74	0.18	0.36	0.54	0.71

For Artificial Data Set 2, λ's were determined by using TPO and BIC criteria, and the most optimal λ value was chosen as $\lambda = 11.03$ according to BIC criteria.

When Table 3.4 is examined, it is seen that PP-RPCA explained 74% of the total variance with the first four main components. When PP-RSPCA results are analyzed, it is seen that 71% of the total variance is explained by the first four main components.

In this case, the cost of using sparsity is equivalent to loss of 3% variance explanation rate. However, in Table 3.4, the number of non-zero loadings in PP-RPCA is 23 while it is seen that the number of non-zero loadings in PP-RSPCA is reduced to 4.

It is seen that the sparsity feature added to the loadings matrix facilitates the interpretation of the principal components with extremely good performance and a low error rate.

3.3. Comparison of PP-RPCA and PP-RSPCA over Artificial Data Set 3 $(p > n)$

The Artificial Data Set 3 has been analyzed with R code written and the results are presented in Table 3.5. For Artificial Data Set 3, λ's were determined by using TPO and BIC criteria, and the most optimal λ value was selected as λ = 50.69 according to BIC criteria.

When Table 3.5 is examined, it is seen that PP-RPCA explains 87% of the total variance with the first four main components. When PP-RSPCA results are analyzed, it is seen that 84% of the total variance is explained by the first four main components. In this case, the cost of using sparsity is equivalent to loss of 3% variance explanation rate. However, it is seen that the number of non-zero loadings in PP-RPCA in Table 3.5 is 71 while the number of non-zero loadings in PP-RSPCA is reduced to 4. It is seen that the sparsity feature added to the loadings matrix facilitates the interpretation of the principal components with extremely good performance and a low error rate.

Table 3.5. The first four non sparse robust (λ = 0) and sparse robust (λ = 50.69) principal component loadings of the variables in Artificial Dataset 3

Variables	PP-RPCA				PP-RSPCA			
	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4
X1	0.62	-0.18	0.36	-0.25	0.00	0.00	0.00	1.00
X2	0.35	0.03	-0.10	0.13	0.00	0.00	0.00	0.00
X3	-0.22	0.25	-0.22	-0.10	0.00	0.00	0.00	0.00
X4	-0.04	-0.15	0.00	-0.02	0.00	0.00	0.00	0.00
X5	-0.24	-0.14	-0.13	0.33	1.00	0.00	0.00	0.00
X6	0.00	0.00	0.04	0.04	0.00	0.00	0.00	0.00

X7	0.00	-0.14	-0.07	0.00	0.00	0.00	0.00	0.00
X8	0.13	0.03	-0.22	-0.07	0.00	0.00	0.00	0.00
X9	0.02	0.50	0.23	0.13	0.00	0.00	0.00	0.00
X10	0.00	-0.13	-0.41	-0.50	0.00	1.00	0.00	0.00
X11	-0.20	0.16	-0.11	-0.08	0.00	0.00	0.00	0.00
X12	-0.08	0.13	0.05	-0.03	0.00	0.00	0.00	0.00
X13	0.11	-0.24	-0.12	0.05	0.00	0.00	0.00	0.00
X14	0.00	0.50	0.07	0.00	0.00	0.00	0.00	0.00
X15	-0.03	-0.21	-0.28	0.55	0.00	0.00	1.00	0.00
X16	0.01	-0.08	-0.07	0.01	0.00	0.00	0.00	0.00
X17	0.00	-0.22	0.14	0.03	0.00	0.00	0.00	0.00
X18	-0.08	0.06	-0.34	-0.41	0.00	0.00	0.00	0.00
X19	0.17	0.25	-0.10	0.07	0.00	0.00	0.00	0.00
X20	-0.53	-0.23	0.51	-0.22	0.00	0.00	0.00	0.00
VER %	0.28	0.24	0.21	0.14	0.27	0.23	0.19	0.16
Cumulative VER %	0.28	0.53	0.73	0.87	0.27	0.50	0.68	0.84

3.4. Comparison of PP-RPCA and PP-RSPCA on Real Data 1- Stock Market Data

Real Data 1-Stock Market Data is analyzed with R code written and the results are presented in Table 3.6.

Table 3.6. The first four non sparse robust ($\lambda = 0$) and sparse robust ($\lambda = 3.31$) principal component loadings of the variables in Real Data 1-Stock Market Data

Variables	PP-RPCA				PP-RSPCA			
	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4
X1	0.36	-0.12	-0.48	0.19	0.00	0.00	0.00	1.00
X2	-0.02	0.56	-0.22	-0.26	0.79	-0.01	0.00	0.00
X3	0.29	0.39	0.19	-0.45	0.00	0.00	0.17	0.01
X4	0.22	0.00	-0.34	0.18	0.00	0.00	0.00	0.00
X5	0.15	0.54	0.11	0.02	0.00	0.00	0.99	0.00
X6	0.59	0.00	-0.35	0.01	0.01	0.10	0.00	0.00
X7	0.46	0.09	0.61	0.55	0.00	1.00	0.00	0.00
X8	-0.39	0.47	-0.26	0.61	0.62	0.01	0.00	0.00

VER %	0.34	0.24	0.15	0.15	0.28	0.19	0.19	0.10
Cumulative VER %	0.34	0.58	0.73	0.87	0.28	0.47	0.66	0.76

For the Stock Market Data, $\boldsymbol{\lambda}$'s were determined by using TPO and BIC criteria, and the most optimal $\boldsymbol{\lambda}$ value was selected as $\boldsymbol{\lambda} = 3.31$ according to TPO criteria.

Table 3.6 shows that PP-RPCA explains 87% of the total variance with the first four main components. When PP-RSPCA results are analyzed, it is seen that 76% of the total variance is explained by the first four main components. In this case, the cost of using the sparsity feature is equivalent to the loss of 11% variance explanation rate. However, in Table 3.6, it is seen that the number of non-zero loadings in PP-RPCA is 30 while the number of non-zero loadings in PP-RSPCA is reduced to 11. It is seen that the sparsity feature added to the loadings matrix facilitates the interpretation of the principal components with extremely good performance and a low error rate.

4. DISCUSSION AND CONCLUSION

The sparse principal component analysis is very useful in finding vectors in the direction of maximizing the variance described and improving the interpretability of the principal components. The robust sparse principal component analysis also allows the identification of principal component analysis vectors, which also facilitate interpretability and are unaffected by outliers.

The projection-pursuit approach discussed maximizes a robust variance to find the principal component analysis vectors. One of the advantages of the projection-pursuit approach is that the principal components are calculated sequentially and the algorithm stops when the desired number of principal components is reached. This feature is particularly useful in cases where the number of variables is greater than the number of observations $(p > n)$ in high-dimensional data sets.

In the study, the performance of the RSPCA method suggested by Croux et al. (2013) was examined on one real and three artificial data sets with different outliers ratios. Artificial dataset 3 includes the number of variables (p) > number of observations (n). Therefore, the findings to be obtained in the solution of such data set with RSPCA are very important. The projection-pursuit method is also a widely used robust method in case of $p > n$.

The optimal value of the sparse parameter λ used to obtain the optimal result in terms of both interpretability and variance explanation rate can be found by BIC and TPO approaches. The BIC and TPO criteria used in the determination of lambda did not show much superiority compared to each other. However, in each application, the optimum of the λ determination criteria (BIC and TPO), which gives the λ value, has been selected.

In the future study, RSPCA based on projection-pursuit discussed in this study can be compared with RSPCA results based on different robust variance estimators. For small sample and large sample cases, the performance of the methods for different outlier ratios can be evaluated. The effects of the use of different approaches such as MCD instead of projection-pursuit approach on RSPCA can also be examined.

REFERENCES

Bien, J., Xu, Y., Mahoney, M. W. 2010. CUR from a sparse optimization viewpoint. *In Advances in Neural Information Processing Systems* (pp. 217-225).

Cadima, J., Jolliffe, I. T. 1995. Loading and correlations in the interpretation of principle components. *Journal of Applied Statistics*, 22(2): 203-214.

Croux, C., Ruiz-Gazen, A. 2005. High breakdown estimators for principal components: the projection-pursuit approach revisited. *Journal of Multivariate Analysis*, 95(1): 206-226.

Croux, C., Filzmoser, P., Oliveira, M. R. 2007. Algorithms for Projection–Pursuit robust principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 87(2): 218-225.

Croux, C., Filzmoser, P., Fritz, H. 2013. Robust sparse principal component analysis. *Technometrics*, 55(2): 202-214.

D'Aspremont, A., El Ghaoui, L., Jordan, M. I., Lanckriet, G. R. 2007. A direct formulation for sparse PCA using semidefinite programming. *SIAM review*, 49(3): 434-448.

Guo, J., James, G., Levina, E., Michailidis, G., Zhu, J. 2010. Principal component analysis with sparse fused loadings. *Journal of Computational and Graphical Statistics*, 19(4): 930-946.

He, Y., Zhou, W., Qian, G., Chen, B. 2014. Methane storage in metal–organic frameworks. *Chemical Society Reviews*, 43(16): 5657-5678.

Hubert, M., Rousseeuw, P. J., Van Aelst, S. 2008. High-breakdown robust multivariate methods. *Statistical Science*, 92-119.

Hubert, M., Reynkens, T., Schmitt, E., Verdonck, T. 2016. Sparse PCA for high-dimensional data with outliers. *Technometrics,* 58(4): 424-434.

Jenatton, R., Obozinski, G., Bach, F. 2010, March. Structured sparse principal component analysis. *In Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics* (pp. 366-373).

Jolliffe, I. T. 1995. Rotation of principal components: choice of normalization constraints. *Journal of Applied Statistics*, 22(1): 29-35.

Jolliffe, I. T., Trendafilov, N. T., Uddin, M. 2003. A modified principal component technique based on the LASSO. *Journal of computational and Graphical Statistics*, 12(3): 531-547.

Journée M., Nesterov Y., Richtárik P., Sepulchre R. 2010 Generalized power method for sparse principal component analysis. *J Mach Learn Res* 11:517–553.

Li, G., Chen, Z. 1985. Projection-pursuit approach to robust dispersion matrices and principal components: primary theory and Monte Carlo. *Journal of the American Statistical Association*, 80(391): 759-766.

Locantore, N., Marron, J. S., Simpson, D. G., Tripoli, N., Zhang, J. T., Cohen, K. L., ..., Fan, J. 1999. Robust principal component analysis for functional data. *Test*, 8(1): 1-73.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rousseeuw, P. J., Croux, C. 1993. Alternatives to the median absolute deviation. *Journal of the American Statistical association*, 88(424): 1273-1283.

Rousseeuw, P. J., Driessen, K. V. 1999. A fast algorithm for the minimum covariance determinant estimator. *Technometrics*, 41(3): 212-223.

Sigg, C. D., Buhmann, J. M. 2008, July. Expectation-maximization for sparse and non-negative PCA. *In Proceedings of the 25th international conference on Machine learning*, ACM, pp. 960-967.

Wang, H., Li, B., Leng, C. 2009. Shrinkage tuning parameter selection with a diverging number of parameters. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(3): 671-683.

Zou, H., Hastie, T., Tibshirani, R. 2006. Sparse principal component analysis. *Journal of computational and graphical statistics*, 15(2): 265-286.

Bilal Barış ALKAN
Department of Educational Measurement and Evaluation,
University of Akdeniz, Antalya 07000, Turkey.
Email: bbarisalkan@gmail.com (corresponding author)

Işıl ÜNALDI
Department of Department of Biostatistics and Medical Informatics,
University of Ondokuz Mayıs, Samsun 55000, Turkey.
Email: isil.unaldi@omu.edu.tr

Solving Lonely Runner Conjecture through differential geometry

V. ĎURIŠ, T. ŠUMNÝ, D. GONDA AND T. LENGYELFALUSY

Abstract

The Lonely Runner Conjecture is a known open problem that was defined by Wills in 1967 and in 1973 also by Cusick independently of Wills. If we suppose n runners having distinct constant speeds start at a common point and run laps on a circular track with a unit length, then for any given runner, there is a time at which the distance of that runner is at least $1/n$ from every other runner. There exist several hypothesis verifications for different n mostly based on principles of approximation using number theory. However, the general solution of the conjecture for any n is still an open problem. In our work we will use a unique approach to verify the Lonely Runner Conjecture by the methods of differential geometry, which presents a non-standard solution, but demonstrates to be a suitable method for solving this type of problems. In the paper we will show also the procedure to build an algorithm that shows the possible existence of a solution for any number of runners.

Mathematics Subject Classification 2000: 53Z05, 11D72
Keywords: lonely runner problem, geometric interpretation, differential geometry, kinematics

1. INTRODUCTION

The conjecture, today known as "Lonely Runner Conjecture," was introduced by Wills (1967) [1] and independently by Cusick (1973) [2]. In its original form, it was formulated as follows: „*For any* n *positive integers* $w_1, w_2, \dots, w_k$, *there is a real number* x *such that*

$$\|w_i x\| \geq \frac{1}{k+1},$$

for each $i = 1, 2, \dots, k$, *where for a real number* x, $\|w_i x\|$ *is the distance of real number* x *from the nearest integer to* x.“.

The name of this hypothesis, "Lonely Runner Conjecture", is the result of an interpretation made by Bienia et al. (1998) [3]: “*At time zero,* k *runners start from the beginning of the circular track with a unit length to run the repeated laps. Each runner maintains a constant non-zero speed. Then there is a time when all runners are at*

10.2478/jamsi-2022-0002

least $1/\ (k+1)$ *away from their common starting point, regardless of their speeds.*”. The term "lonely runner" refers to an equivalent formulation in which there are $k+1$ runners with different speeds. Then there is the time in which the runner is "lonely", that is, at a distance of at least $1/(k+1)$ from the others.

The proof of Lonely Runner Conjecture for $k=2$ is very simple. For $k=3$ there are more approaches. Wills (1967) [1], for example, considered this problem in terms of Diophantine approximation, independently of Cusick (1973) [2]. Cusick considered a k-dimensional "view-obstruction" problem. The case of $k=4$ was first proved by Cusick and Pomerance (1984) [4], and this proof required computer control. Later, Bienia et al. (1998) [3] provided simpler proof for the case $k=4$. Case $n=5$ was proved by Bohman, Holzman and Kleitman (2001) [5]. A simpler proof for case $n=5$ was provided by Renault (2004) [6]. Barajas and Serra (2008) [7] proved the hypothesis for $n=6$ and also solve the case of the conjecture for $n=7$ [8]. The work [8] focuses on a specific case for seven runners, using congruences for dividing the track into appropriate intervals. Pandey (2009) [9] proved the conjecture for two or more runners provided the speed of the $(i+1)$-th runner is more than double the speed of the i-th runner for each i, with the speeds arranged in an increasing order. Finding a universal conjecture verification for any n is still an open problem. Some authors have verified the hypothesis for selected large n when determining initial assumptions, e.g. in Dubickas (2011) [10] we can find conjecture verification for $n>16342$ under assumption that the speeds of the runners satisfy $\frac{v_{j+1}}{v_j}\geq 1+\frac{33\log n}{n}$ for $j=1,\dots,n-1$. The work [10] points to the possibility of creating some distribution intervals based on the ratio of the speeds of neighboring runners. Although this approach is approximate, it can be extended to an unlimited number of runners.

2. MAIN RESULTS

Let there be n runners on a circular track. If each runner runs at a different speed and we let them run for a sufficiently long time, we examine if there occurs a situation where the runners will be evenly distributed on the track. This problem is designed so that the computational methods lead to advances in the field of Diophantine approximation [11].

We will solve the problem using kinematics, by the methods of differential geometry to determine distances. This approach makes it easier for us to create a moving coordinate system. Given that this concept will follow the physical interpretation of the problem, therefore the techniques we will use are based on theoretical physics [12]. For a length element in N-dimensional space with the metric g_{kl}, it holds

$$dl^2 = g_{kl} dq_k dq_l,$$

where we use the generalized coordinates q_k and Einstein's summation rule [13-14]. In addition, we determine the metric based on the inner product of the respective base vectors

$$g_{kl} = e_k \cdot e_l.$$

Consider a non-stationary coordinate system (the analogy of the Galilean transformation [15]) and define a distance for individual length elements

$$dl = \sum_{\substack{k=1 \\ l=1}}^{n} dl_{kl} = (v_k - q_l)dt.$$

Next, consider an oriented space, in which the distance throughout the whole space can be determined as

$$dl^2 = g_{jj} dq_j dq_j = 0,$$

$$dl_{kj} = \left(O_k - v_j dt\right),$$

$$O_k = v_k dt.$$

This means that each runner is considered to be the origin of his coordinate system

$$g_{jj} = 0$$

and

$$g_{kl} = -g_{lk}.$$

The whole track consists of individual partial sections, where the partial sections are determined as the distance between the runner and his nearest co-runner. Thus, when we count all the partial sections, we get the whole circle and we actually return to the position of the first runner.

Let's show a solution for the well-known trivial case of two runners

$$dl = \sum_{\substack{k=1 \\ l=1}}^{n} dl_{kl} = |v_k - q_l| dt.$$

From the point of view of the second runner, we consider that the origin of the coordinate system is at first runner. At time t, the second runner is far from the first runner so that they are exactly opposite each other, then $l = \frac{1}{2}$. By substitution we get

$$\frac{1}{2} = (v_2 - v_1) \cdot t.$$

By simplification we get a known formula for the situation of two runners

$$t = \frac{1}{2 \cdot (v_2 - v_1)}.$$

For the case of several runners, we're looking for the time $\tau = k_i \cdot t_i, k_i \in \mathbb{R}$, which ensues as a shift based on the performed circuits. To illustrate, we can state the problem for three runners, while on the left side of the equation we have the distance determined the way that the two runners are $\frac{1}{3}$ of a distance apart and are also divided by the distance created by the mutual circulation with respect to the starting position. For this reason consider $\alpha \in \mathbb{R}$ and $k_i = n_i + \alpha$ under the condition $n_i \in \mathbb{N}$, where n_i is the number of circuits and α is the shift of the runner closest to the starting position from the starting position. Hence we can determine the system of equations

$$\frac{1}{3} + k_1 = (v_2 - v_1) \cdot t,$$
$$\frac{1}{3} + k_2 = (v_3 - v_2) \cdot t,$$
$$\frac{1}{3} + k_3 = (v_3 - v_1) \cdot t.$$

After adding the equations we get

$$1 + K = 2(v_3 - v_1)t,$$

while $K = k_1 + k_2 + k_2 + 3\alpha$. Then

$$t = \frac{1+K}{2(v_3 - v_1)}.$$

After substituting t we have

$$\frac{1}{3} + k_3 = \frac{1+k_1+k_2+k_2}{2}.$$

From that after subsequent simplification it holds

$$k_3 = \frac{1}{3} + k_1 + k_2.$$

In this way we can sequentially obtain next values for k_1 or k_2. This way we get one open parameter and a formula for the other parameters. The resulting relationship between the two parameters is very similar to the linear Diophantine equation, so even here we can search the solution using lattice points [16]. The geometric representation of the solution is a line passing through the grid points, and we can determine the solvability by the parameter $a \in \mathbb{R}$, which also determines the shift of the runners from the starting position. In addition, we see that the Diophantine equation provides us with a whole set of solutions.

Finally, we will show how to solve the case for n runners. By applying a length element it holds

$$dl = \sum_{\substack{k=1 \\ l=1}}^{n} dl_{kl} = |v_k - q_l| dt.$$

From the point of view of individual runners with respect to the nearest neighbor, we get a system of equations in the form

$$\frac{1}{n} + k_i = |v_i - v_{i+1}| \cdot t,$$

where $k_i = n_i + \alpha$ under the condition $n_i \in \mathbb{N}$, $\alpha \in \mathbb{R}$. Due to the fact that we have a cyclical situation, the time will depend mostly on the runners we placed as first and last, specifically v_1 and v_n. Then we are able to determine the resulting formula for the time needed

$$t = \frac{1+\sum k_i}{2(v_3 - v_1)}.$$

Hence, by sequentially substitution we get the Diophantine equation for k_i and k_j and open parameter $\alpha \in \mathbb{R}$, which provides the solvability. This approach is equivalent to the consequences we would achieve by applying the Kantorovich metric applied to a circle, which is shown in detail in the work [17]. In this article, the authors solve the resulting measures on a circle based on the division of the circle. Although the considered measure is applied to statistics, its fundamentals is to create a measure that in equity divides the circle. This is an analogy to the distribution of runners in our approach.

3. CONCLUSION

The paper focused on the use of differential geometry methods in solving the problem of the "Lonely Runner Conjecture". Although most methods are based on the principles of approximation, we focused on the geometric interpretation of the problem. We inserted a circle with a unit length into the coordinate system in such a way that each runner created his own coordinate system. Based on this, we created a mapping that describes the distribution of runners from the perspective of any runner. Given that our approach is close to the solution of a physical problem and thus also the analysis of a problem as kinematics problem, we've used this method for solving lonely runner problem. We also pointed out that the solution matches the known solution for two runners and subsequently we applied it to a situation with three runners. The provided solution gives prescription for an algorithm that increases its difficulty with the growing number of runners. On the other hand, it shows the existence of a solution for any number of runners, what was the main contribution of our article.

REFERENCES

[1] Wills J. M. (1967). *Zwei Sätze über inhomogene diophantische Approximation von Irrationalzehlen*. In. Monatshefte für Mathematik, Vol. 71, No. 3, pp. 263-269.

[2] Cusick T. W. (1973). *View-obstruction problems*. In. Aequationes mathematicae, Vol. 9, No. 2, pp. 165-170, DOI: 10.1007/BF01832623.

[3] Bienia W., Goddyn L., Gvozdjak P., Sebő A., Tarsi M. (1998). *Flows, view obstructions, and the lonely runner*. In. Journal of combinatorial theory, Series B, Vol. 72, No. 1, pp. 1-9, DOI: 10.1006/jctb.1997.1770.

[4] Cusick T. W., Pomerance C. (1984). *View-obstruction problems, III*. In. Journal of Number Theory, Vol. 19, No. 2, pp. 131-139, DOI: 10.1016/0022-314X(84)90097-0.

[5] Bohman T., Holzman R., Kleitman D. (2001). *Six lonely runners*. In. Electronic Journal of Combinatorics, Vol. 8, No. 2, pp. 1-49, DOI: 10.37236/1602.

[6] Renault J. (2004). *View-obstruction: a shorter proof for 6 lonely runners*. In. Discrete mathematics, Vol. 287, No. 1-3, pp. 93-101, DOI 10.1016/j.disc.2004.06.008.

[7] Barajas J., Serra O. (2008). *Distance graphs with maximum chromatic number*. In. Discrete mathematics, Vol. 308, No. 8, pp. 1355-1365, DOI 10.1016/j.disc.2007.07.061.

[8] Barajas J., Serra O. (2008). *The lonely runner with seven runners*. In. The electronic journal of combinatorics, Vol. 15, pp. 1-18, DOI: 10.37236/772.

[9] Pandey R. K. (2009). *A note on the lonely runner conjecture*. In. Journal of Integer Sequences, Vol. 12, pp. 1-4.

[10] Dubickas A. (2011). *The lonely runner problem for many runners*. In. Glasnik Matematički, Vol 46, No 1, pp. 25-30, DOI: 10.3336/gm.46.1.05.

[11] Lang S. (1995). *Introduction to Diophantine Approximations: New Expanded Edition*. New York: Spronger Verlag, 2nd ed., 140 p., ISBN: 9780387944562.

[12] Kulhánek P. (2016). *Selected chapters from theoretical physics*. Prague: Aldebaran, Language: Czech, 416 p., ISBN: 9788090458284.

[13] Einstein A. (1916). *The Special and General Theory*. 100 p., London: Methuen Publishing Ltd.

[14] Tullis J. (2020). *Einstein Summation Crash Crash Course: An Introduction to Index and Tensor Notation*. USA: 50-50 Publishing LLC, 39 p., ISBN: 9798565340396.

[15] Yaglom I. M. (1979). *A Simple Non-Euclidean Geometry and Its Physical Basis: An Elementary Account of Galilean Geometry and the Galilean Principle of Relativity*. New York: Springer-Verlag, 307 p., ISBN: 9780387903323.

[16] Ďuriš V., Šumný T. (2019). *Diophantine Geometry in Space E2 and E3*. In. TEM Journal, ISSN 2217-8309, Vol. 8, No. 1, pp. 78-81, DOI 10.18421/TEM81-10.

[17] Cabrelli C. A., Molter U. M. (1995). *The Kantorovich metric for probability measures on the circle*. In. Journal of Computational and Applied Mathematics, Vol. 57, No. 3, pp. 345-361, DOI: 10.1016/0377-0427(93)E0213-6.

Viliam Ďuriš
Constantine The Philosopher University in Nitra
Tr. A. Hlinku 1, 94901 Nitra, Slovakia
Email: vduris@ukf.sk

Timotej Šumný
Štefan Moyses Primary School
Školská 608, Tesárske Mlyňany, Slovakia
Email: sumnyt@gmail.com

Dalibor Gonda
University of Žilina
Univerzitná 1, 01001 Žilina, Slovakia
Email: daliborgonda@gmail.com

Tomáš Lengyelfalusy
DTI University
Sládkovičova 533/20, 018 41 Dubnica nad Váhom, Slovakia
Email: lengyelfalusy@dti.sk

A new ratio type estimator for computation of population mean under post-stratification

K. UL ISLAM RATHER, M. IQBAL JEELANI, M. YOUNIS SHAH, S. E. H. RIZVI AND M. SHARMA

Abstract

In this study, the difficulty of estimating the population mean in the situation of post-stratification is discussed. The case of post-stratification is presented for ratio-type exponential estimators of finite population mean. Mean-squared error of the proposed estimator is obtained up to the first degree of approximation. In the instance of post-stratification, the proposed estimator was compared with the existing estimators. An empirical study by using some real data and further, simulation study has been carried out to demonstrate the performance of the proposed estimator.

Mathematics Subject Classification 2000: 62D05
Keywords: post-stratification, auxiliary variable, exponential ratio estimator, mean squared error, efficiency.

1. INTRODUCTION

One of the most prevalent survey sample approaches is stratification. The use of stratified random sampling ensures that the size of each stratum as well as the sampling frame for each stratum is known ahead of time. However, in many practical cases, the overall population size and percentage of the unit that falls into distinct strata or stratum sizes are known, but a sampling frame for each stratum is either not accessible or is too expensive and time-consuming to produce. We can't use stratified random sampling in these scenarios. To overcome this problem, a post-stratification procedure is utilized, in which a sample of the requisite size is taken from the population using SRS and then stratified using the stratification factor. Many authors contributed in the field of post stratification including Holt and Smith (1979), Jagers et al. (1985), Jagers (1986), Ige and Tripathi (1989), Agrawal and Panda (1993), and Singh and Ruiz Espezo (2003).

We assume a finite population of size N which is divided into L strata of sizes $N_1, N_2, \ldots, N_L$ such that $\sum_{h=1}^{L} N_h = N$. Let y be the study variate, and x be the

10.2478/jamsi-2022-0003

auxiliary variate, positively and negatively correlated with study variate y, respectively. Let y_{hi} be the observation on i^{th} unit of h^{th} stratum for study variate, and x_{hi} be the observation on i^{th} unit of h^{th} stratum for auxiliary variate x. For study variate y and auxiliary variate x population means are denoted by $\bar{Y}$ and $\bar{X}$, respectively, while h^{th} stratum means are denoted by $\bar{Y}_h$ and $\bar{X}_h$ respectively. A sample of size n is drawn from the population using simple random sampling without replacement (SRSWR). It is observed that how many and which units belong to the h^{th} stratum. Let n_h be the size of the sample falling in, h^{th} stratum such that $\sum_{h=1}^{L} n_h = n$ here, it is assumed that n is so large that possibility of n_h being zero is very small.

After stratification, ratio and product estimators for population mean were discussed Ige and Tripathi (1989). Tailor investigated the ratio exponential estimator of Bahl and Tuteja (1991) in post-stratification Tailor *et al.* (2017). The authors are motivated to research ratio and ratio type exponential estimators in post-stratification as a result of their previous work. In post stratification, usual unbiased estimator for population mean is explained as.

$$\bar{y}_{ps} = \sum_{h=1}^{H} W_h \bar{y}_h \tag{1.1}$$

where, $\frac{N_h}{N}$ is the weight of h^{th} stratum and $\bar{y}_h$ is the sample mean of n_h sample units that fall in the h^{th} stratum.

Using the results from Stephen (1945), the variance is given as

$$Var\left(\hat{\bar{Y}}_{ps}\right) = \left(\frac{1}{n} - \frac{1}{N}\right) \sum_{h=1}^{L} W_h\, S_{yh}^2 + \frac{1}{n^2} \sum_{h=1}^{L} (1 - W_h)\, S_{yh}^2 \tag{1.2}$$

Ige and Tripathi (1989) defined classical ratio type estimator in post-stratification, as

$$\hat{\bar{Y}}_R = \bar{y}_{ps} \left(\frac{\bar{X}}{\bar{x}_{ps}} \right) \tag{1.3}$$

where

$$\bar{X} = \sum_{h=1}^{L} W_h \bar{X}_h \; and \quad \bar{y}_{ps} = \sum_{h=1}^{L} W_h \bar{y}_h, \qquad \bar{x}_{ps} = \sum_{h=1}^{L} W_h \bar{x}_h$$

Upto the fda, the bias and MSE statement of $\hat{\bar{Y}}_{ps}^{R}$ is defined as

$$B\left(\hat{\bar{Y}}_R\right) = \left(\frac{1}{n} - \frac{1}{N}\right)\frac{1}{\bar{X}}\sum_{h=1}^{L} W_h\left(R_1 S_{xh}^2 - S_{yxh}\right) \tag{1.4}$$

And

$$MSE\left(\hat{\bar{Y}}_R\right) = \left(\frac{1}{n} - \frac{1}{N}\right)\sum_{h=1}^{L} W_h\left(S_{yh}^2 + R_1^2 - 2R_1 S_{yxh}\right) \tag{1.5}$$

where $R_1 = \frac{\bar{Y}}{\bar{X}}$

Tailor *et al.* (2017) defined Bahl and Tuteja (1991) ratio type exponential estimator as

$$\hat{\bar{Y}}_{RC} = \bar{y}_{ps} exp\left(\frac{\bar{x}_{ps} - \bar{X}}{\bar{x}_{ps} + \bar{X}}\right) \tag{1.6}$$

Up to the fda, the bias and MSE assertion of $\hat{\bar{Y}}_{RC}$ are respectively given by

$$B\left(\hat{\bar{Y}}_{RC}\right) = \left(\frac{1}{n} - \frac{1}{N}\right)\frac{1}{\bar{X}}\sum_{h=1}^{L} W_h\left(\frac{3}{8} R_1 S_{xh}^2 + \frac{1}{2} S_{yxh}\right) \tag{1.7}$$

$$MSE\left(\hat{\bar{Y}}_{RC}\right) = \left(\frac{1}{n} - \frac{1}{N}\right)\sum_{h=1}^{L} W_h\left(S_{yh}^2 + \frac{1}{4} R_1^2 S_{xh}^2 - R_1 S_{yxh}\right) \tag{1.8}$$

$$\hat{\bar{Y}}_{PC} = \bar{y}_{ps} exp\left(\frac{\bar{X} - \bar{x}_{ps}}{\bar{X} + \bar{x}_{ps}}\right) \tag{1.9}$$

Up to the fda, the bias and MSE assertion of $\hat{\bar{Y}}_R$ are respectively given by

$$B\left(\hat{\bar{Y}}_{PC}\right) = \left(\frac{1}{n} - \frac{1}{N}\right)\frac{1}{\bar{\bar{X}}}\sum_{h=1}^{L} W_h\left(\frac{3}{8}R_1 S_{xh}^2 - \frac{1}{2}S_{yxh}\right) \tag{1.10}$$

$$MSE\left(\hat{\bar{Y}}_{PC}\right) = \left(\frac{1}{n} - \frac{1}{N}\right)\sum_{h=1}^{L} W_h\left(S_{yh}^2 + \frac{1}{4}R_1^2 S_{xh}^2 - R_1 S_{yxh}\right) \tag{1.11}$$

2. PROPOSED ESTIMATOR

Motivated by Subramani (2016), we have proposed the following general ratio type estimator of population mean of study variable under post stratification.

$$\hat{\bar{Y}}_{RK} = \bar{y}_{PS}\left[k + (1-k)exp\left(\frac{\bar{X} - \bar{x}_{ps}}{\bar{X} + \bar{x}_{ps}}\right)\right] \tag{2.1}$$

where k is constant and we write,

To obtain the MSE of $\hat{\bar{Y}}_{RK}$, write

$e_0 = \frac{1}{\bar{\bar{Y}}}\sum_{h=1}^{L} W_h \bar{Y}_h$ and $e_1 = \frac{1}{\bar{\bar{X}}}\sum_{h=1}^{L} W_h \bar{X}_h$

such that

$E(e_0) = E(e_1) = 0$, and

$$E(e_0^2) = \frac{1}{\bar{\bar{Y}}^2}\left(\frac{1}{n} - \frac{1}{N}\right)\sum_{h=1}^{l} W_h\, S_{yh}^2$$

$$E(e_1^2) = \frac{1}{\bar{\bar{X}}^2}\left(\frac{1}{n} - \frac{1}{N}\right)\sum_{h=1}^{L} W_h\, S_{xh}^2,$$

$$E(e_0 e_1) = \frac{1}{\bar{\bar{Y}}\bar{\bar{X}}}\left(\frac{1}{n} - \frac{1}{N}\right)\sum_{h=1}^{L} W_h\, S_{yxh}$$

The proposed estimator $\hat{\hat{\mathrm{Y}}}_{\mathrm{RK}}$ expressed in terms of e_i,s

$$\hat{\hat{Y}}_{RK} = \bar{Y}(1+e_0)\left\{k+(1-k)exp\left(\frac{-\bar{X}e_1}{2\bar{X}+\bar{X}e_1}\right)\right\}$$

$$\hat{\hat{Y}}_{RK} = \bar{Y}(1+e_0)\left\{k+(1-k)exp\left(\frac{-e_{1h}}{(2+e_{1h})}\right)\right\}$$

$$\hat{\hat{Y}}_{RK} = \bar{Y}(1+e_0)\left\{k+(1-k)exp\left(\frac{-e_1}{2}\left(1+\frac{e_1}{2}\right)^{-1}\right)\right\} \tag{2.2}$$

Expanding the right hand side of (2.2) and retaining terms up to the second power of *e*’s,

$$\hat{\hat{Y}}_{RK} = \bar{Y}(1+e_0)\left\{k+(1-k)exp\left(\frac{-e_1}{2}-\frac{e_1^2}{4}\right)\right\}$$

$$MSE\left(\hat{\hat{Y}}_{RK}\right) = \left\{E(e_0^2)+k^2E\left(\frac{e_1^2}{4}\right)-kE(e_0e_1)\right\}$$

$$MSE\left(\hat{\hat{Y}}_{RK}\right) = f\sum_{h=1}^{L}W_h\left(S_{yh}^2+\frac{1}{4}k^2R_1^2S_{xh}^2-kR_1S_{yxh}\right) \tag{2.3}$$

2.1. Optimality of κ

Obtain the optimum *k* to minimize $MSE\left(\hat{\hat{Y}}_{RK}\right)$. Differentiating $MSE\left(\hat{\hat{Y}}_{RK}\right)$ with respect to *k* and equating the derivative to zero, optimum value of *k* is given by

$$k^* = 2\sum_{h=1}^{L}\frac{W_hS_{yxh}}{W_hR_1S_{xh}^2} \tag{2.4}$$

Substituting the value of k^* in (2.3), we get the minimum value of $MSE\left(\hat{\hat{Y}}_{RK(min)}\right)$

$$MSE\left(\hat{\hat{Y}}_{RK(min)}\right) = f\sum_{h=1}^{L}\left[S_{yh}^2+2\left(\frac{W_hS_{yxh}^2}{W_hS_{xh}^2}\right)\right] \tag{2.5}$$

3. EFFICIENCY COMPARISON

In this section, the performance of the proposed estimator has been demonstrated over the existing estimators in the literature as follows:

From (1.2) and (2.5)

I. $\left[Var\left(\hat{\hat{Y}}_{ps}\right) - MSE\left(\hat{\hat{Y}}_{RK(min)}\right)\right] > 0$

$$\frac{1}{n^2}\sum_{h=1}^{L}(1-W_h)\,S_{yh}^2 - \sum_{h=1}^{L}W_h\left(\frac{2S_{yxh}^2}{S_{xh}^2}\right) > 0 \tag{3.1}$$

From (1.5) and (2.5)

II. $\left[MSE\left(\hat{\hat{Y}}_{R}\right) - MSE\left(\hat{\hat{Y}}_{RK(min)}\right)\right] > 0$

$$\sum_{h=1}^{L}W_h\left(R_1^2 - 2R_1S_{yxh}\right) - \sum_{h=1}^{L}W_h\left(\frac{2S_{yxh}^2}{S_{xh}^2}\right) > 0 \tag{3.2}$$

From (1.8) and (2.5)

III. $\left[MSE\left(\hat{\hat{Y}}_{\mathrm{RC}}\right) - \mathrm{MSE}\left(\hat{\hat{Y}}_{\mathrm{RK(min)}}\right)\right] > 0$

$$\sum_{h=1}^{L}W_h\left(\frac{1}{4}R_1^2S_{xh}^2 - R_1S_{yxh}\right) - \sum_{h=1}^{L}W_h\left(\frac{2S_{yxh}^2}{S_{xh}^2}\right) > 0 \tag{3.3}$$

From (1.11) and (2.5)

IV. $\left[MSE\left(\hat{\hat{Y}}_{PC}\right) - MSE\left(\hat{\hat{Y}}_{RK(min)}\right)\right] > 0$

$$\sum_{h=1}^{L}W_h\left(\frac{1}{4}R_1^2S_{xh}^2 + R_1S_{yxh}\right) - \sum_{h=1}^{L}W_h\left(\frac{2S_{yxh}^2}{S_{xh}^2}\right) > 0 \tag{3.4}$$

It is observed that $\hat{\hat{Y}}_{RK(min)}$ is always more efficient than the traditional estimators $\hat{\hat{Y}}_{ps}$, $\hat{\hat{Y}}_{R}$, $\hat{\hat{Y}}_{\mathrm{RC}}$ and $\hat{\hat{Y}}_{PC}$, because the conditions from (3.1) to (3.4) are always satisfied.

4. EMPIRICAL STUDY

A data set is considered to exhibits the performance of the suggested estimator; we use some real-life populations. Description of the populations is given below:

Population I [Source: Murthy (1967), p. 228]

y : Output,

x: Fixed capital

Constant	Stratum I	Stratum II
N_h	05	05
n_h	4	4
$\bar{Y}_h$	1925	3115.60
$\bar{X}_h$	214.40	333.80
S_{yh}	615.92	340.38
S_{xh}	74.87	66.35
S_{yxh}	39360.68	22356.50

Population II [Source: Murthy (1967), p. 228]

y : Output,

x: Fixed capital

Constant	Stratum I	Stratum II
N_h	05	05
n_h	2	2
$\bar{Y}_h$	1925	3115.60
$\bar{X}_h$	214.40	333.80
S_{yh}	615.92	340.38
S_{xh}	74.87	66.35
S_{yxh}	39360.68	22356.50

Population III

[Source: Japan Meteorological Society. Retrieved from the World Wide Web http://www.data.jma.go.jp/obd/stats/data/en/index.html**]**

Constant	**Stratum I**	**Stratum II**
N_h	10	10
n_h	4	4
$\bar{Y}_h$	1629.99	2035.96
$\bar{X}_h$	149.7	102.6
S_{yh}	32.305	103.26
S_{xh}	74.872	66.35
S_{yxh}	-1072.8	-655.25

Table 4.1 Mean square errors (MSEs) of the estimators for Population I:

Estimators	**MSE**
$V(\bar{y}_{ps})$	10059.07
$\hat{\bar{Y}}_R$	2580.72
$\hat{\bar{Y}}_{RC}$	15929.74
$\hat{\bar{Y}}_{PC}$	1740.528
$\boldsymbol{\hat{\bar{Y}}^*_{RK}}$	**1432.67**

Table 4.2 PRE of Estimators for Population I:

Estimators	**PRE**
$V(\bar{y}_{ps})$	100
$\hat{\bar{Y}}_R$	389.78
$\hat{\bar{Y}}_{RC}$	63.15
$\hat{\bar{Y}}_{PC}$	577.93
$\boldsymbol{\hat{\bar{Y}}^*_{RK}}$	**702.12**

Table 4.2, revealed the percent relative efficiencies (PRE) of estimators for population I. It is observed that the proposed exponential ratio type estimator $\bar{Y}_{RK}$ proved to be the best estimator in the sense of having highest percent relative efficiency than usual unbiased estimators$V(\bar{y}_{ps})$, $\hat{\bar{Y}}_R$, $\hat{\bar{Y}}_{RC}$, $\hat{\bar{Y}}_{PC}$ for the population I.

Table 4.3 Mean square errors (MSEs) of the estimators for Population II:

Estimators	MSE
$V(\bar{y}_{ps})$	52617.18
$\hat{\bar{Y}}_R$	15483.09
$\hat{\bar{Y}}_{RC}$	95578.39
$\hat{\bar{Y}}_{PC}$	10443.09
$\boldsymbol{\hat{\bar{Y}}^*_{RK}}$	**8595.62**

Table 4.4 PRE of Estimators for Population II:

Estimators	PRE
$V(\bar{y}_{ps})$	100
$\hat{\bar{Y}}_R$	339.84
$\hat{\bar{Y}}_{RC}$	55.05
$\hat{\bar{Y}}_{PC}$	503.85
$\boldsymbol{\hat{\bar{Y}}^*_{RK}}$	**612.14**

Table 4.4, revealed the percent relative efficiencies (PRE) of estimators for population II. It is observed that the proposed exponential ratio type estimator $\bar{Y}_{RK}$ proved to be the best estimator in the sense of having highest percent relative

efficiency than usual unbiased estimators $V(\bar{y}_{ps}), \hat{\bar{Y}}_R$, $\hat{\bar{Y}}_{RC}$, $\hat{\bar{Y}}_{PC}$ for the population II.

Table 4.5 Mean square errors (MSEs) of the estimators for Population III:

Estimators	MSE
$V(\bar{y}_{ps})$	530.48
$\hat{\bar{Y}}_R$	2709.97
$\hat{\bar{Y}}_{RC}$	112.85
$\hat{\bar{Y}}_{PC}$	1304.72
$\boldsymbol{\hat{\bar{Y}}^*_{RK}}$	**109.90**

Table 4.6 PRE of Estimators for Population III:

Estimators	PRE
$V(\bar{y}_{ps})$	100
$\hat{\bar{Y}}_R$	19.58
$\hat{\bar{Y}}_{RC}$	470.06
$\hat{\bar{Y}}_{PC}$	40.66
$\boldsymbol{\hat{\bar{Y}}^*_{RK}}$	**482.67**

Table 4.4, revealed the percent relative efficiencies (PRE) of estimators for population III. It is observed that the proposed exponential ratio type estimator $\bar{Y}_{RK}$ proved to be the best estimator in the sense of having highest percent relative efficiency than usual unbiased estimators $V(\bar{y}_{ps}), \hat{\bar{Y}}_R$, $\hat{\bar{Y}}_{RC}$, $\hat{\bar{Y}}_{PC}$ for the population III.

5. SIMULATION STUDY

We did a simulation research in this section to explore the qualities of proposed estimators. We use a bivariate normal distribution to produce finite populations of size N=10000 in the simulation study. The samples were created from a bivariate normal distribution using the R program's mvrnorm function. In the simulation, we considered $\mu_x = 2, \mu_y = 4,$ we have computed mean square errors (MSEs) and percent relative efficiencies (PREs) of estimators with respect to $V(\bar{y}_{ps})$ and displayed in Table 5.1- Table 5.2. From the Tables we can say that proposed estimator $\hat{\bar{Y}}_{RK}^{*}$ is the most efficient estimator than the existing estimators in literature for this simulation study.

Table 5.1 MSE and PRE Values of Estimators for population I

$\rho_{yx1} : \rho_{yx2}$	*MSE*	*PRE*
	10059.07	100
	4249.29	236.72
0.75:0.88	15095.45	66.64
	2574.82	390.67
	2551.59	**294.23**
	10059.07	100
	3459.59	290.76
0.80:0.93	15490.3	64.94
	7179.96	461.43
	2022.02	**497.48**
	10059.07	100
	2669.88	376.76
0.85:0.98	15885.16	63.32
	1785.11	563.50
	1492.46	**673.99**

Table 5.2 MSE and PRE Values of Estimators for population II

$\rho_{yx1} : \rho_{yx2}$	*MSE*	*PRE*
0.75:0.88	52617.17	100
	25494.61	206.39
	90572.64	58.09
	15448.85	340.59
	15309.35	**343.69**
0.80:0.93	52617.17	100
	20756.34	253.50
	92941.77	56.61
	13079.72	402.28
	12131.87	**431.71**
0.85:0.98	52617.17	100
	16018.07	318.49
	95310.91	55.206
	10710.59	491.26
	8954.38	**587.61**

6. CONCLUSION

In this article, an exponential ratio type estimator has been proposed under post stratification. The mathematical form of the estimator has been derived and its condition of efficiencies has been formulated with respect to some existing estimators from literature. For comparing the efficiencies of proposed estimator with some existing estimators, we utilized some real-life populations for estimating population mean. The results from these populations show that our proposed estimator performs efficiently as compared to existing estimators. We also observe that efficiency of proposed estimator increases when the correlation between study and auxiliary variable increases by simulation study. Therefore, it is recommended to use proposed estimator for estimating population mean under post stratification.

ACKNOWLEDGMENT

Authors are thankful to the referee for his valuable suggestions regarding the improvement of the paper.

REFERENCES

Agrawal, M. C., Panda, K. B. (1993). An efficient estimator in post stratification. *Metron* 51:179–188.

Bahl, S., Tuteja, R. K. (1991). Ratio and product type exponential estimator. *Inf. Optim. Sci.* 12:159–163.

Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). Wiley Eastern Limited.

Holt, D., Smith, T. M. F. (1979). Post-stratification. *J. R. Stat. Soc.* 142:33–46.

Ige, A. F., Tripathi, T. P. (1989). Estimation of population mean using post-stratification and auxiliary information. *Abacus* 18:265–276.

Jagers, P. (1986). Post stratification against bias in sampling. *Int. Statist. Rev.* 55:159–167.

Jagers, P., Oden, A., Trulsson, L. (1985). Post stratification and ratio estimation: Usages of auxiliary information in survey sampling and opinion polls. *Int. Statist. Rev.* 53:221–238.

Murthy, M. N. (1967). *Sampling Theory and Methods* (p. 228). Calcutta: Statistical Publishing Society.

Singh, H. P., Ruiz Espejo, M. (2003). *Improved post stratified estimation*. Bulletin of the International Statistical Institute, 54th session, contributed papers, vol. LX, Two Books, Book 2, pp. 341–342.

Singh, R., Mangat, N. S . (1996). *Elements of Survey Sampling* (Vol. , p. 219). London: Kluwer Academic.

Stephan, F. (1945). The expected value and variance of the reciprocal and other negative powers of a positive Bernoullian variate. *Ann. Math. Stat.* 16:50–61.

Subramani, J. (2016). A new median based ratio estimator for estimation of the finite population mean, Statistics in Transition New Series, 17, 4:1-14.

Tailor R., Tailor R., and Chouhan S. (2017). Improved Ratio- and Product-Type Exponential Estimators for Population Mean in Case of Post-Stratification. *Communications in Statistics - Theory and Methods,* Vol 46, No. 21, 10387-10393.

Tailor R., Mehta P. (2019). A Ratio and Ratio Exponential Estimator for Finite Population Mean in Case of Post-Stratification. *Journal of Statistics Applications & Probability.* 8(3):241-246.

Khalid Ul Islam Rather
Division of Statistics and Computer Science, Faculty of Basic Sciences,
Main Campus SKUAST-J, Jammu, J&K, India, 180009.
Email: khalidstat34@gmail.com (corresponding author)

M. Iqbal Jeelani
Division of Statistics and Computer Science, Faculty of Basic Sciences,
Main Campus SKUAST-J, Jammu, J&K, India, 180009.

M. Younis Shah
Division of Statistics and Computer Science, Faculty of Basic Sciences,
Main Campus SKUAST-J, Jammu, J&K, India, 180009.

S. E. H. Rizvi
Division of Statistics and Computer Science, Faculty of Basic Sciences,
Main Campus SKUAST-J, Jammu, J&K, India, 180009.

Manish Sharma
Division of Statistics and Computer Science, Faculty of Basic Sciences,
Main Campus SKUAST-J, Jammu, J&K, India, 180009.

Asymptotic expectation of protected node profile in random digital search trees

M. JAVANIAN, R. IMANY NABIYYI, J. TOOFANPOUR AND M. Q. VAHIDI-ASL

Abstract

Protected nodes are neither leaves nor parents of any leaves in a rooted tree. We study here protected node profile, namely, the number of protected nodes with the same distance from the root in digital search trees, some fundamental data structures to store 0 - 1 strings. When each string is a sequence of independent and identically distributed Bernoulli(p) random variables with $0 < p < 1$ ($p \neq \frac{1}{2}$), Drmota and Szpankowski (2011) investigated the expectation of internal profile by the analytic methods. Here, we generalize the main parts of their approach in order to obtain the asymptotic expectations of protected node profile and non-protected node profile in digital search trees.

Mathematics Subject Classification 2000: 05C05, 68Q87, 35A20
Keywords: Digital search trees, profile, protected node, singularity analysis, saddle point method, probability generating function.

1. INTRODUCTION

Digital search trees (DSTs) are one of the most essential varieties of data structures for strings in computer science algorithms (see [12] for more information).

A DST is a binary tree built from a sequence of 0 - 1 strings. The first string is stored in the root node. The subsequent string is directed to the left subtree if its first digit is 0, or to the right if the first digit is 1. This process works recursively for the subtrees of the root but, at level i, the $(i+1)$th digit of the string is used for branching. Internal nodes hold the strings. See Figure 1 for an example of a DST built from 20 internal nodes.

We assume here that each string is a sequence of independent and identically distributed Bernoulli(p) random variables with $0 < p < 1$, the probability of occurring a “1”; we also use $0 < q := 1 - p < p < 1$. A DST built from this sequence of Bernoulli random variables is referred to as a random DST. Random DSTs have been extensively studied by many authors, see e.g. [2], [3], [4], [11], [14], [15], and the references therein.

10.2478/jamsi-2022-0004

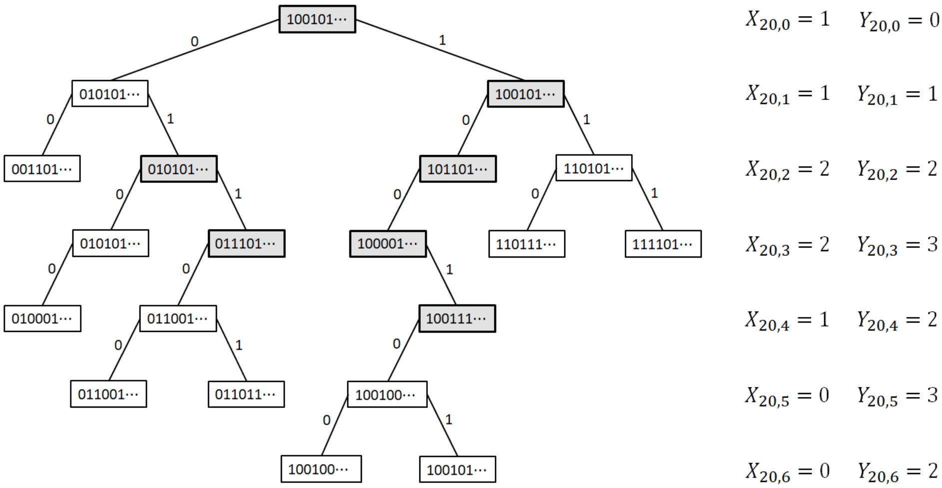

Fig. 1. A DST built on 20 strings stored in internal nodes (rectangles) with protected (gray ones) and non-protected (white ones) nodes, and its profiles.

By profiles, the most important shape parameters in rooted trees, we mean the number of nodes of the same type at the same level. These parameters are connected to some other shape parameters such as height, width. See [3], [4], [9], [10], [11], [17], [16] for different types of results on profiles in some classes of random trees.

A node is said to be protected if it is not a leaf and, furthermore, none of its children is a leaf; otherwise a node is said to be non-protected. *E.g.*, Figure 1 shows the protected and non-protected nodes in a DST. Protected nodes have been studied for many different random trees in recent papers; see for instance, [1], [13], [7] and the papers cited there.

In this paper, we investigate the protected node profile, namely, the number of protected nodes with the same distance from the root in random DSTs. Throughout the paper, we write $I_{n,k}$, $X_{n,k}$ and $Y_{n,k}$ for, respectively, the number of internal nodes, the number of protected nodes and the number of non-protected internal nodes at level k in a DST of size n (e.g. see Figure 1 for the values of $X_{n,k}$ and $Y_{n,k}$ in a DST). These profiles have been analyzed by [9] for tries, and by [10] for recursive trees, newly. For DSTs, as $p = \frac{1}{2}$ and $n \to \infty$, the mean, the variance and the asymptotic normality of $I_{n,k}$ are fully clarified in [4]. As $p \neq \frac{1}{2}$, the asymptotics of the mean and variance of $I_{n,k}$ have been obtained in [3] and [11], respectively. Here, in the case of $p \neq \frac{1}{2}$, we generalize the main part of the proof in [3] to obtain the asymptotic means of $X_{n,k}$ and $Y_{n,k}$.

2. SETTING

In a random DST of size n, the number of protected nodes $X_{n,k}$ at level $k \geq 1$, can be computed recursively by computing the number for the two subtrees at level $k-1$. For $k=0$, the root is protected, if and only if neither the left nor the right subtree contains only one string. This leads to the following distributional recurrence for $X_{n,k}$:

$$X_{n,k} \stackrel{d}{=} \begin{cases} X_{B_n,k-1} + X^*_{n-1-B_n,k-1}, & k \geq 1; \\ 1 - \mathbf{1}_{\{1,n-2\}}(B_n), & k = 0, \end{cases} \qquad n \geq 2,$$

where $\stackrel{d}{=}$ denotes equality in distribution, $B_n \stackrel{d}{=} \text{Binomial}(n-1,p)$, $X_{n,k} \stackrel{d}{=} X^*_{n,k}$, $\mathbf{1}_A(\cdot)$ is the indicator function of A and $X_{n,k}$, $X^*_{n,k}$, B_n are independent. Moreover, $X_{n,k} = 0$ for $n \leq 1$ and $k \geq 0$. Similarly, if $Y^*_{n,k}$ is distributed as $Y_{n,k}$ then we have

$$Y_{n,k} \stackrel{d}{=} \begin{cases} Y_{B_n,k-1} + Y^*_{n-1-B_n,k-1}, & k \geq 1; \\ \mathbf{1}_{\{1,n-2\}}(B_n), & k = 0, \end{cases} \qquad n \geq 2,$$

where $Y_{n,k}$, $Y^*_{n,k}$, B_n are independent. Furthermore, the initial boundary conditions are $Y_{0,k} = 0$ for $k \geq 0$, $Y_{1,0} = 1$, $Y_{1,k} = 0$ for $k \geq 1$.

Let $\varphi_{n,k}(u,w) := \mathbb{E}[u^{X_{n,k}} w^{Y_{n,k}}]$ be the joint probability generating function of $X_{n,k}$ and $Y_{n,k}$. Then, by the above recurrences, we obtain

$$\varphi_{n,k}(u,w) = \sum_{j=0}^{n-1} \binom{n-1}{j} p^j q^{n-1-j} \varphi_{j,k-1}(u,w) \varphi_{n-1-j,k-1}(u,w), \quad n \geq 1, k \geq 1, \tag{1}$$

with the initial and boundary conditions

$$\varphi_{0,k}(u,w) = 1, \quad k \geq 0, \qquad \varphi_{1,0}(u,w) = w,$$
$$\varphi_{n,0}(u,w) = u + (n-1)(pq^{n-2} + p^{n-2}q)(w-u), \quad n \geq 2.$$

Denote the Poisson generating function of $\varphi_{n,k}(u,w)$ by

$$\tilde{\varphi}_k(z,u,w) := e^{-z} \sum_{n \geq 0} \varphi_{n,k}(u,w) \frac{z^n}{n!}, \qquad \tilde{\varphi}_k(0,u,w) = \varphi_{0,k}(u,w) = 1, \quad k \geq 0.$$

This definition and the relation (1) fulfills the following recurrence

$$\frac{\partial \tilde{\varphi}_k(z,u,w)}{\partial z} + \tilde{\varphi}_k(z,u,w) = \tilde{\varphi}_{k-1}(pz,u,w) \tilde{\varphi}_{k-1}(qz,u,w), \quad k \geq 1, \tag{2}$$

$$\begin{aligned} \tilde{\varphi}_0(z,u,w) &= e^{-z} + wze^{-z} + u(1 - e^{-z} - ze^{-z}) \\ &\quad + (w-u)\left(\frac{p}{q^2}(qz-1)e^{-pz} + \frac{q}{p^2}(pz-1)e^{-qz} + \frac{p}{q^2}e^{-z} + \frac{q}{p^2}e^{-z}\right) \end{aligned} \tag{3}$$

From (2) and (3), the following Poisson transforms

$$\tilde{M}_k^{[X]}(z) := e^{-z} \sum_{n \geq 0} \mathbb{E}(X_{n,k}) \frac{z^n}{n!} = \frac{\partial \tilde{\varphi}_k(z,u,w)}{\partial u}\Big|_{u=w=1},$$

$$\tilde{M}_k^{[Y]}(z) := e^{-z} \sum_{n \geq 0} \mathbb{E}(Y_{n,k}) \frac{z^n}{n!} = \frac{\partial \tilde{\varphi}_k(z,u,w)}{\partial w}\Big|_{u=w=1},$$

satisfy the recurrence relations, for $k \geq 1$,

$$\tilde{M}_k'^{[X]}(z) + \tilde{M}_k^{[X]}(z) = \tilde{M}_{k-1}^{[X]}(pz) + \tilde{M}_{k-1}^{[X]}(qz), \tag{4}$$

$$\tilde{M}_k'^{[Y]}(z) + \tilde{M}_k^{[Y]}(z) = \tilde{M}_{k-1}^{[Y]}(pz) + \tilde{M}_{k-1}^{[Y]}(qz), \tag{5}$$

with

$$\tilde{M}_0^{[X]}(z) = 1 - e^{-z} - ze^{-z} - \frac{p}{q^2}(qz-1)e^{-pz} - \frac{q}{p^2}(pz-1)e^{-qz} - \Big(\frac{p}{q^2} + \frac{q}{p^2}\Big)e^{-z},$$

$$\tilde{M}_0^{[Y]}(z) = ze^{-z} + \frac{p}{q^2}(qz-1)e^{-pz} + \frac{q}{p^2}(pz-1)e^{-qz} + \Big(\frac{p}{q^2} + \frac{q}{p^2}\Big)e^{-z}.$$

Since $\tilde{M}_0^{[I]}(z) = 1 - e^{-z}$ has been obtained in [3], then the above initial conditions confirm that

$$\tilde{M}_0^{[I]}(z) = 1 - e^{-z} = \tilde{M}_0^{[X]}(z) + \tilde{M}_0^{[Y]}(z). \tag{6}$$

By induction it is easy to prove that $\tilde{M}_k^{[Y]}(z)$ can be represented as a finite linear combination of the two functions of the forms $e^{-p^{l_1}q^{l_2}z}$ and $ze^{-p^{l_1}q^{l_2}z}$ with $l_1, l_2 \geq 0$ and $l_1 + l_2 \leq k+1$. Furthermore, $X_{n,k} = 0$ for $k > n-3$; and $Y_{n,k} = 0$ for $k > n-1$. Thus, $\tilde{M}_k^{[X]}(z) = \mathscr{O}(z^{k+3})$ and $\tilde{M}_k^{[Y]}(z) = \mathscr{O}(z^{k+1})$ for $z \to 0$ which ensures that $M_k^{*[X]}(s)$ exists for s with $-k-3 < \Re(s) < 0$; and $M_k^{*[Y]}(s)$ exists for s with $\Re(s) > -k-1$. Let us now express $M_k^{*[X]}(s)$ and $M_k^{*[Y]}(s)$ as

$$M_k^{*[X]}(s) = -\Gamma(s)F_k^{[X]}(s), \qquad M_k^{*[Y]}(s) = -\Gamma(s)F_k^{[Y]}(s)$$

where $\Gamma(s)$ is the Euler gamma function. From the above setting, $F_k^{[X]}(s)$ and $F_k^{[Y]}(s)$ are some finite linear combinations of functions of the forms $p^{-l_1 s}q^{-l_2 s}$ and $sp^{-l_1 s}q^{-l_2 s}$ with $l_1, l_2 \geq 0$ and $0 \leq l_1 + l_2 \leq k+1$. By (4) and (5),

$$\begin{aligned} F_k^{[X]}(s) - F_k^{[X]}(s-1) &= T(s)F_{k-1}^{[X]}(s), \\ F_k^{[Y]}(s) - F_k^{[Y]}(s-1) &= T(s)F_{k-1}^{[Y]}(s), \end{aligned} \qquad k \geq 1,$$

with $T(s) := p^{-s} + q^{-s}$ and the initial conditions

$$F_0^{[X]}(s) = 1 + s + \frac{p^{-s}}{q^2}(qs - p) + \frac{q^{-s}}{p^2}(ps - q) + \frac{p}{q^2} + \frac{q}{p^2}$$
$$F_0^{[Y]}(s) = -s - \frac{p^{-s}}{q^2}(qs - p) - \frac{q^{-s}}{p^2}(ps - q) - \frac{p}{q^2} - \frac{q}{p^2}.$$

3. THE EXPECTATIONS

Here, in order to obtain the expectations of the protected and non-protected profiles, i.e. $\mathbb{E}(X_{n,k})$ and $\mathbb{E}(Y_{n,k})$, we first derive the asymptotic expansions of $F_k^{[X]}(s)$ and $F_k^{[Y]}(s)$ (these two functions are defined in the previous section), by generalizing the analysis of $F_k^{[I]}(s) := M_k^{*[I]}(s)/\Gamma(s)$ ($M_k^{*[I]}(s)$ is the Mellin transform of $\tilde{M}_k^{[I]}(z)$) used to derive the expectation of internal profile $\mathbb{E}(I_{n,k})$ in [3]. Then we invert the Mellin transforms $M_k^{*[X]}(s)$ and $M_k^{*[Y]}(s)$. Finally, we invert the Poisson transforms $\tilde{M}_k^{[X]}(z)$ and $\tilde{M}_k^{[Y]}(z)$ to obtain the asymptotics for the expected profiles, $\mathbb{E}(X_{n,k})$ and $\mathbb{E}(Y_{n,k})$.

In the lemma below, we first find some explicit representations of $F_k^{[X]}(s)$ and $F_k^{[Y]}(s)$.

LEMMA 3.1. The functions $F_k^{[X]}(s)$ and $F_k^{[Y]}(s)$ are recursively given by

$$\begin{aligned} F_k^{[X]}(s) &= \mathbf{A}[F_{k-1}^{[X]}](s) - \mathbf{A}[F_{k-1}^{[X]}](-1), \\ F_k^{[Y]}(s) &= \mathbf{A}[F_{k-1}^{[Y]}](s) - \mathbf{A}[F_{k-1}^{[Y]}](-1), \end{aligned} \qquad k \geq 1,$$

where the operator $\mathbf{A}$ is defined by $\mathbf{A}[f](s) = \sum_{j\geq 0} f(s-j)T(s-j)$; $T(s) = p^{-s} + q^{-s}$,

$$F_0^{[X]}(s) = 1 + s + \frac{p^{-s}}{q^2}(qs - p) + \frac{q^{-s}}{p^2}(ps - q) + \frac{p}{q^2} + \frac{q}{p^2} =: 1 - F(s),$$
$$F_0^{[Y]}(s) = -s - \frac{p^{-s}}{q^2}(qs - p) - \frac{q^{-s}}{p^2}(ps - q) - \frac{p}{q^2} - \frac{q}{p^2} =: F(s),$$

$$\begin{aligned} F_k^{[X]}(-r) &= 0, \qquad r = 1, 2, \ldots, k+2, \\ F_k^{[Y]}(-r) &= 0, \qquad r = 0, 1, 2, \ldots, k. \end{aligned} \qquad k \geq 1,$$

Furthermore, if we set $R_k(s) := \mathbf{A}^k[1](s)$ and $\widehat{R}_k(s) := \mathbf{A}^k[F](s)$ then we have

$$\begin{aligned} \sum_{k\geq 0} F_k^{[X]}(s) w^k &= -\sum_{k\geq 0} \widehat{R}_k(s) w^k + \sum_{k\geq 0} \widehat{R}_k(-1) w^k \cdot \frac{\sum_{k\geq 0} R_k(s) w^k}{\sum_{k\geq 0} R_k(-1) w^k}, \\ \sum_{k\geq 0} F_k^{[Y]}(s) w^k &= \sum_{k\geq 0} \widehat{R}_k(s) w^k + \Big(1 - \sum_{k\geq 0} \widehat{R}_k(-1) w^k\Big) \frac{\sum_{k\geq 0} R_k(s) w^k}{\sum_{k\geq 0} R_k(-1) w^k}. \end{aligned} \tag{7}$$

PROOF. Set $\tilde{F}_0^{[X]}(s)=F_0^{[X]}(s)$, $\tilde{F}_0^{[Y]}(s)=F_0^{[Y]}(s)$ and recursively

$$\begin{aligned}\tilde{F}_k^{[X]}(s)&=\mathbf{A}[\tilde{F}_{k-1}^{[X]}](s)-\mathbf{A}[\tilde{F}_{k-1}^{[X]}](-1),\\ \tilde{F}_k^{[Y]}(s)&=\mathbf{A}[\tilde{F}_{k-1}^{[Y]}](s)-\mathbf{A}[\tilde{F}_{k-1}^{[Y]}](-1),\end{aligned}\qquad k\geq 1.$$

We first show that $\tilde{F}_k^{[X]}(s)=F_k^{[X]}(s)$ and $\tilde{F}_k^{[Y]}(s)=F_k^{[Y]}(s)$ by induction on k, namely, the same argument as the first part of the proof of Theorem 3 in [3]: $\tilde{F}_k^{[X]}(s)$ is (as it is for $F_k^{[X]}(s)$) a finite linear combinations of functions of the forms $p^{-l_1 s}q^{-l_2 s}$ and $sp^{-l_1 s}q^{-l_2 s}$ with $l_1,l_2\geq 0$ and $l_1+l_2\leq k+1$. By definitions, we have $\tilde{F}_k^{[X]}(-1)=0$,

$$\tilde{F}_0^{[X]}(s)=F_0^{[X]}(s)\quad\text{and}\quad \tilde{F}_k^{[X]}(s)-\tilde{F}_k^{[X]}(s-1)=T(s)\tilde{F}_{k-1}^{[X]}(s),\qquad k\geq 1.$$

Now suppose $\tilde{F}_k^{[X]}(s)=F_k^{[X]}(s)$ holds for some $k\geq 0$. Then $\tilde{F}_{k+1}^{[X]}(s)=F_{k+1}^{[X]}(s)+G^{[X]}(s)$, where $G^{[X]}(s)$ satisfies $G^{[X]}(s)-G^{[X]}(s-1)=0$. It follows that $G^{[X]}(s)$ is a finite linear combinations of functions of the forms $p^{-l_1 s}q^{-l_2 s}$. Since $G^{[X]}(s)$ is a periodic function such that $G^{[X]}(s)=G^{[X]}(s-1)$ then $G^{[X]}(s)$ is a zero function. Hence $\tilde{F}_{k+1}^{[X]}(s)=F_{k+1}^{[X]}(s)$. Similarly we can conclude that $\tilde{F}_k^{[Y]}(s)=F_k^{[Y]}(s)$.

Since $1/\Gamma(-r)=0$ for $r=0,1,2,\ldots,k+2$ and $\tilde{F}_k^{[X]}(s)=M_k^{*[X]}(s)/\Gamma(s)$ is analytic for $-k-3<\Re(s)<0$ then $F_k^{[X]}(-r)=0$ for $r=1,\ldots,k+2$. Similarly we have $\tilde{F}_k^{[Y]}(-r)=0$ for $r=0,1,2,\ldots,k$ ($\tilde{F}_k^{[Y]}(s)$ is analytic for $\Re(s)>-k-1$).

Now we prove (7). Let the generating functions of $F_k^{[X]}(s)$ and $F_k^{[Y]}(s)$ be defined by

$$f^{[X]}(s,w):=\sum_{k\geq 0}F_k^{[X]}(s)w^k\quad\text{and}\quad f^{[Y]}(s,w):=\sum_{k\geq 0}F_k^{[Y]}(s)w^k$$

with $f^{[X]}(-1,w)=0$ and $f^{[Y]}(-1,w)=1$. If $\mathbf{I}$ be the identity matrix then we have

$$\begin{aligned}f^{[X]}(s,w)&=F_0^{[X]}(s)+w\sum_{k\geq 1}\Big(\mathbf{A}[F_{k-1}^{[X]}](s)-\mathbf{A}[F_{k-1}^{[X]}](-1)\Big)w^{k-1}\\&=F_0^{[X]}(s)+w\mathbf{A}\big[f^{[X]}(\cdot,w)\big](s)-w\mathbf{A}\big[f^{[X]}(\cdot,w)\big](-1),\end{aligned}$$

which is equivalent to

$$(\mathbf{I}-w\mathbf{A})\big[f^{[X]}(\cdot,w)\big](s)=F_0^{[X]}(s)-w\mathbf{A}\big[f^{[X]}(\cdot,w)\big](-1),$$

or

$$\begin{aligned}f^{[X]}(s,w)&=(\mathbf{I}-w\mathbf{A})^{-1}[F_0^{[X]}](s)-\Big(w\mathbf{A}\big[f^{[X]}(\cdot,w)\big](-1)\Big)(\mathbf{I}-w\mathbf{A})^{-1}[1](s)\\&=-\sum_{k\geq 0}\widehat{R}_k(s)w^k+\Big(1-w\mathbf{A}\big[f^{[X]}(\cdot,w)\big](-1)\Big)\sum_{k\geq 0}R_k(s)w^k.\end{aligned}$$

Therefore we derive

$$f^{[X]}(s,w) = -\sum_{k\geq 0}\widehat{R}_k(s)w^k + \Big(f^{[X]}(-1,w) + \sum_{k\geq 0}\widehat{R}_k(-1)w^k\Big)\frac{\sum_{k\geq 0}R_k(s)w^k}{\sum_{k\geq 0}R_k(-1)w^k}.$$

Finally, we obtain the similar result for $f^{[Y]}(s,w)$ and then the claims in (7).

REMARK 3.2. Using $F_k^{[X]}(-r) = 0$ for $k > r-3$ and $F_k^{[Y]}(-r) = 0$ for $k > r-1$, and setting $s = -r$ in (7), we find

$$\begin{aligned}
\sum_{k=0}^{r-3} F_k^{[X]}(-r)w^k &= -\sum_{k\geq 0}\widehat{R}_k(-r)w^k + \sum_{k\geq 0}\widehat{R}_k(-1)w^k\cdot\frac{\sum_{k\geq 0}R_k(-r)w^k}{\sum_{k\geq 0}R_k(-1)w^k},\\
\sum_{k=0}^{r-1} F_k^{[Y]}(-r)w^k &= \sum_{k\geq 0}\widehat{R}_k(-r)w^k + \Big(1-\sum_{k\geq 0}\widehat{R}_k(-1)w^k\Big)\frac{\sum_{k\geq 0}R_k(-r)w^k}{\sum_{k\geq 0}R_k(-1)w^k},
\end{aligned}\tag{8}$$

and consequently

$$\begin{aligned}
f^{[X]}(s,w) &= -\sum_{k\geq 0}\widehat{R}_k(s)w^k + \Big(\sum_{k=0}^{r-3}F_k^{[X]}(-r)w^k + \sum_{k\geq 0}\widehat{R}_k(-r)w^k\Big)\frac{\sum_{k\geq 0}R_k(s)w^k}{\sum_{k\geq 0}R_k(-r)w^k},\\
f^{[Y]}(s,w) &= \sum_{k\geq 0}\widehat{R}_k(s)w^k + \Big(\sum_{k=0}^{r-1}F_k^{[Y]}(-r)w^k - \sum_{k\geq 0}\widehat{R}_k(-r)w^k\Big)\frac{\sum_{k\geq 0}R_k(s)w^k}{\sum_{k\geq 0}R_k(-r)w^k}.
\end{aligned}\tag{9}$$

In the lemma bellow, we will analyze the generating function $\widehat{g}(s,w) := \sum_{k\geq 0}\widehat{R}_k(s)w^k$.

LEMMA 3.3. There exists a function $\widehat{h}(s,w)$ that is analytic for all w and s satisfying

$$wT(s-m)\neq 1, \quad \text{for all } m\geq 1,$$

such that, for $F(s) = -s - q^{-2}p^{-s}(qs-p) - p^{-2}q^{-s}(ps-q) - pq^{-2} - p^{-2}q$, we have

$$\widehat{g}(s,w) := \sum_{k\geq 0}\widehat{R}_k(s)w^k := \sum_{k\geq 0}\mathbf{A}^k[F](s)w^k = \frac{\widehat{h}(s,w)}{1-wT(s)},\tag{10}$$

where the operator $\mathbf{A}$ is defined by $\mathbf{A}[f](s) = \sum_{j\geq 0} f(s-j)T(s-j)$; $T(s) = p^{-s}+q^{-s}$. Thus, $\widehat{g}(s,w)$ has a meromorphic continuation where $w_0 = 1/T(s)$ is a polar singularity.

PROOF. Since $\widehat{R}_{k+1}(s) = \mathbf{A}[\widehat{R}_k](s)$ and $\widehat{R}_0(s) = F(s)$, then, for $p > q$, we have

$$\begin{aligned}
|\widehat{R}_{k+1}(s)| &\leq \sum_{j\geq 0}|\widehat{R}_k(s-j)|\cdot p^j\cdot T(\Re(s)), \qquad k\geq 0,\\
|\widehat{R}_0(s)| &\leq \frac{|s|}{q}T(\Re(s)) + \frac{p}{q^2}T(\Re(s)) + \frac{2|s|}{q} + \frac{2p}{q^2}\\
&\leq \begin{cases} 2|s|q^{-1}T(\Re(s)) + 2pq^{-2}T(\Re(s)), & \text{if } \Re(s)\geq 0;\\ 4|s|q^{-1} + 4pq^{-2}, & \text{if } \Re(s) < 0.\end{cases}
\end{aligned}\tag{11}$$

Therefore, it follows by induction on k that

$$|\widehat{R}_k(s)| \leq \begin{cases} 2\prod_{j\geq 1}(1-p^j)^{-2}\big(|s|+\frac{1}{1-p}\big)T(\Re(s))^{k+1}, & \text{if } \Re(s)\geq 0;\\ 4\prod_{j\geq 1}(1-p^j)^{-2}\big(|s|+\frac{1}{1-p}\big)T(\Re(s))^{k}, & \text{if } \Re(s)<0; \end{cases} \qquad k\geq 0.$$

Thus, if $|w| < T(\Re(s))^{-1}$ the series

$$\widehat{g}(s,w) = \sum_{k\geq 0}\widehat{R}_k(s)w^k = \Big(\sum_{k\geq 0} w^k\mathbf{A}^k\Big)[F](s) = (\mathbf{I}-w\mathbf{A})^{-1}[F](s)$$

converges absolutely and represents an analytic function. Then $\widehat{g}(s,w)$ satisfies

$$\widehat{g}(s,w) - w\sum_{j\geq 0}\widehat{g}(s-j,w)T(s-j) = (\mathbf{I}-w\mathbf{A})[\widehat{g}(\cdot,w)](s) = F(s). \tag{12}$$

Rewrite (12) by substituting $\frac{\widehat{h}(s,w)}{1-wT(s)}$ in for $\widehat{g}(s,w)$. So we obtain

$$\widehat{h}(s,w) = F(s) + \sum_{j\geq 1}\widehat{h}(s-j,w)\frac{wT(s-j)}{1-wT(s-j)}. \tag{13}$$

Recall that $\widehat{h}(s,w) = \widehat{g}(s,w)\big(1-wT(s)\big)$ exists for $|w| < T(\Re(s))^{-1}$. Now, we will use (13) to show that $\widehat{h}(s,w)$ can be analytically continued to all w such that $wT(s-m)\neq 1$, for all $m\geq 1$.

In order to show this claim, we define the following operator $\mathbf{B}$ and function $U(s,w)$,

$$\mathbf{B}[f](s) := \sum_{j\geq 1} f(s-j,w)\frac{wT(s-j)}{1-wT(s-j)}, \qquad U(s,w) := \frac{wT(s)}{1-wT(s)}.$$

By induction it follows that

$$\begin{aligned}\mathbf{B}^k[F](s) &= \sum_{i_1\geq 1}\cdots\sum_{i_k\geq 1} U(s-i_1,w)\cdots U(s-i_1-\cdots-i_k,w)F(s-i_1-\cdots-i_k,w)\\ &= \sum_{m_k\geq k}\sum_{m_{k-1}=k-1}^{m_k-1}\cdots\sum_{m_1=1}^{m_2-1} U(s-m_1,w)\cdots U(s-m_{k-1},w)U(s-m_k,w)F(s-m_k,w).\end{aligned}$$

Therefore,

$$\begin{aligned}\big|\mathbf{B}^k[F](s)\big| &\leq \sum_{m_k\geq k}\cdots\sum_{m_1\geq 1}\big|U(s-m_1,w)\cdots U(s-m_k,w)F(s-m_k,w)\big|\\ &\leq \sum_{m_1\geq 1}\big|U(s-m_1,w)\big|\cdots\sum_{m_k\geq k}\big|U(s-m_k,w)\big|\big|F(s-m_k,w)\big|.\end{aligned}$$

From (11) and $T(s-m)=\mathcal{O}(q^m)$, we find out that the following two series

$$\sum_{m\geq 1}\big|U(s-m,w)\big|\big|F(s-m,w)\big| = \sum_{m\geq 1}\frac{|wT(s-m)||F(s-m_k,w)|}{|1-wT(s-m)|},$$
$$S:=\sum_{m\geq 1}\big|U(s-m,w)\big| = \sum_{m\geq 1}\frac{|wT(s-m)|}{|1-wT(s-m)|},$$

converge if $wT(s-m)\neq 1$, for all $m\geq 1$. So there exists an integer $k_0\geq 1$ such that

$$\sum_{m\geq k_0}\big|U(s-m,w)\big|\big|F(s-m,w)\big|\leq \frac{2}{3}, \qquad \sum_{m\geq k_0}\big|U(s-m,w)\big|\leq \frac{2}{3}.$$

Then, for all $k\geq k_0$, we have

$$\big|\mathbf{B}^k[F](s)\big|\leq S^{k_0}\big(\frac{2}{3}\big)^{k-k_0}.$$

Furthermore, $\sum_{k\geq 0}\big|\mathbf{B}^k[F](s)\big|\leq 3\big(\frac{3}{2}S\big)^{k_0}$. In view of this, $\widehat{h}(s,w):=\sum_{k\geq 0}\mathbf{B}^k[F](s)$ is well defined and it satisfies (13).

By Lemma 3 in [3], there exists a function $h(s,w)$ that is analytic for all w and s satisfying $wT(s-m)\neq 1$, for all $m\geq 1$, such that

$$g(s,w):=\sum_{k\geq 0}R_k(s)w^k=:\sum_{k\geq 0}\mathbf{A}^k[1](s)w^k=\frac{h(s,w)}{1-wT(s)} \tag{14}$$

where the operator $\mathbf{A}$ is defined by $\mathbf{A}[f](s)=\sum_{j\geq 0}f(s-j)T(s-j)$; $T(s)=p^{-s}+q^{-s}$.

Now, as bellow, we derive some asymptotic expansions for $F_k^{[X]}(s)$ and $F_k^{[Y]}(s)$.

LEMMA 3.4. Let the operator $\mathbf{B}$ be $\mathbf{B}[f](s)=\sum_{j\geq 1}f(s-j,w)\frac{wT(s-j)}{1-wT(s-j)}$. Set

$$h(s,w)=\sum_{k\geq 0}\mathbf{B}^k[1](s), \qquad \widehat{h}(s,w)=\sum_{k\geq 0}\mathbf{B}^k[F](s),$$

where $F(s)=-s-q^{-2}p^{-s}(qs-p)-p^{-2}q^{-s}(ps-q)-pq^{-2}-p^{-2}q$; $T(s)=p^{-s}+q^{-s}$.

For every real interval $[a,b]$ there exist k_0, $\eta>0$ and $\varepsilon>0$ such that

$$\begin{aligned} F_k^{[X]}(s)&=f^{[X]}(s)T(s)^k\big(1+\mathcal{O}(e^{-\eta k})\big),\\ F_k^{[Y]}(s)&=f^{[Y]}(s)T(s)^k\big(1+\mathcal{O}(e^{-\eta k})\big), \end{aligned} \tag{15}$$

uniformly for all s such that $\Re(s)\in[a,b]$, $|\Im(s)-2l\pi/\log(q/p)|\leq\varepsilon$ for some integer l

and $k \geq k_0$, where, for some $r \in \mathbb{Z}^+$, $f^{[X]}(s)$ and $f^{[Y]}(s)$ are the analytic functions

$$
\begin{aligned}
f^{[X]}(s) &:= \sum_{l=0}^{r-3} F_l^{[X]}(-r)T(s)^{-l}\frac{h(s,1/T(s))}{h(-r,1/T(s))}\Big(1-\frac{T(-r)}{T(s)}\Big)\\
&\quad + \frac{\widehat{h}(-r,1/T(s))h(s,1/T(s))}{h(-r,1/T(s))} - \widehat{h}(s,1/T(s)),\\
f^{[Y]}(s) &:= \sum_{l=0}^{r-1} F_l^{[Y]}(-r)T(s)^{-l}\frac{h(s,1/T(s))}{h(-r,1/T(s))}\Big(1-\frac{T(-r)}{T(s)}\Big)\\
&\quad - \frac{\widehat{h}(-r,1/T(s))h(s,1/T(s))}{h(-r,1/T(s))} + \widehat{h}(s,1/T(s)).
\end{aligned}
\tag{16}
$$

Furthermore, if $|\Im(s) - 2l\pi/\log(q/p)| > \varepsilon$ for all integers l then, for $\Re(s) \in [a,b]$,

$$
F_k^{[X]}(s) = \mathscr{O}\big(T(\Re(s))^k e^{-\eta k}\big), \qquad F_k^{[Y]}(s) = \mathscr{O}\big(T(\Re(s))^k e^{-\eta k}\big). \tag{17}
$$

PROOF. Assume first that $s > -r-1$ for some integer $r \geq 0$ but s is not a positive integer. Using (14), (10) and the representations (9), we have

$$
\begin{aligned}
f^{[X]}(s,w) &= \sum_{k=0}^{r-3} F_k^{[X]}(-r)w^k \frac{h(s,w)}{h(-r,w)}\frac{1-wT(-r)}{1-wT(s)}\\
&\quad - \frac{\widehat{h}(s,w)}{1-wT(s)} + \frac{\widehat{h}(-r,w)}{h(-r,w)}\frac{h(s,w)}{1-wT(s)},
\end{aligned}
\tag{18}
$$

$$
\begin{aligned}
f^{[Y]}(s,w) &= \sum_{k=0}^{r-1} F_k^{[Y]}(-r)w^k \frac{h(s,w)}{h(-r,w)}\frac{1-wT(-r)}{1-wT(s)}\\
&\quad + \frac{\widehat{h}(s,w)}{1-wT(s)} - \frac{\widehat{h}(-r,w)}{h(-r,w)}\frac{h(s,w)}{1-wT(s)}.
\end{aligned}
\tag{19}
$$

From Lemma 3.1 and Lemma 3 in [3], it follows, respectively, that the functions $\widehat{h}(s,w)$ and $h(s,w)$ are analytic for $|w| < 1/T(s-1)$. Hence, $w_0 = 1/T(s)$ is a singularity of $f^{[X]}(s,w)$ and $f^{[Y]}(s,w)$. Namely, the fraction $\widehat{h}(s,w)/(1-wT(s))$, the first summand in both (18) and (19), has no other singularities w with $|w| < 1/T(s)$. By the proof of Lemma 4 in [3], the second summand and, therefore, the third summand in both (18) and (19) have no singularities on the radius of convergence $|w| = 1/T(s)$.

Let $s > -r-1$ is real (but not an integer). Hence, for some $\eta > 0$, by applying Cauchy's formula for a contour of integration on the circle $\gamma = \{w \in \mathbb{C} : |w| = e^{\eta}/T(s)\}$ and the residue theorem [6; 18] we have

$$
F_k^{[X]}(s) = \frac{1}{2\pi i}\int_\gamma f^{[X]}(s,w)w^{-k-1}dw = f^{[X]}(s)T(s)^k + \mathscr{O}\big(|T(s)e^{-\eta}|^k\big),
$$

where $f^{[X]}(s)$ is given in (16). These estimates are uniform for $s \in [a,b] \subseteq (-r-1,-r)$ $(r \geq 0)$ or for $s \in [a,b] \subseteq \mathbb{R}^+$. Furthermore, we obtain the same result if $a \leq \Re(s) \leq b$ and

$|\Im(s)| \le \varepsilon$ for enough small $\varepsilon > 0$.

Next, assume that $s \in \mathbb{R}$ (or s is sufficiently close to the real axis) and close to an integer $-r < 0$, say $s \in [-r-\eta, -r+\eta]$ (for some $\eta > 0$). In this case we have (note that $\Gamma(s)$ is singular at $s = -r$)

$$\begin{aligned}\Gamma(s) \sum_{k>r-3} F_k^{[X]}(s) w^k = &\frac{\Gamma(s)\big(\widehat{h}(-r,w) - \widehat{h}(s,w)\big)}{1 - wT(s)} + \frac{\widehat{h}(-r,w)}{h(-r,w)} \frac{\Gamma(s)\big(h(s,w) - h(-r,w)\big)}{1 - wT(s)} \\ &+ \sum_{k=0}^{r-3} F_k^{[X]}(-r) w^k \frac{\Gamma(s)\big(h(s,w) - h(-r,w)\big)}{h(-r,w)} \frac{1 - wT(-r)}{1 - wT(s)} \\ &+ \sum_{k=0}^{r-3} F_k^{[X]}(-r) w^{k+1} \frac{\Gamma(s)\big(T(s) - T(-r)\big)}{1 - wT(s)}. \end{aligned} \tag{20}$$

Since the functions $\widehat{h}(-r,w)/h(-r,w)$ and $\sum_{k=0}^{r-3} F_k^{[X]}(-r)/h(-r,w)$ are analytic for $|w| < 1/T(-r-1)$ and the point $w = 1/T(s)$ is a polar, then, by (20) and using Cauchy's formula, for $k > r-3$, we obtain $F_k^{[X]}(s)$ in (15).

For integers l, we have $|T(s + 2\pi il/\log(q/p))| = |T(s)|$ due to the fact that

$$T(s + 2\pi il/\log(q/p)) = e^{2\pi il \log(p)/\log(q/p)} T(s).$$

Consequently, $w = 1/T(s)$ is a polar of $f^{[X]}(s,w)$ if $|\Im(s) - 2l\pi/\log(q/p)| < \varepsilon$ for some integer l. Thus, for s in this range, we obtain the estimate of $F_k^{[X]}(s)$ in (15).

Finally, if $|\Im(s) - 2l\pi/\log(q/p)| > \varepsilon$ for all integer l then $|T(s)| < e^{-2\eta}|T(\Re(s))|$ for some $\eta > 0$. Therefore, $f^{[X]}(s,w)$ is regular for $|w| < e^{\eta}/T(\Re(s))$. In conclusion, if we use Cauchy's formula for the contour of integration $\gamma = \{w \in \mathbb{C} : |w| = e^{\eta}/T(\Re(s))\}$, then (17) follows. Here, $\Re(s)$ can vary in a finite interval $[a,b]$.

Similarly, we can prove the estimates for $F_k^{[Y]}(s)$ as given in (15) and (17).

By the above discussion, we know that $F_k^{[X]}(s)$, $F_k^{[Y]}(s)$, $M_k^{*[X]}(s) = -\Gamma(s)F_k^{[X]}(s)$ and $M_k^{*[Y]}(s) = -\Gamma(s)F_k^{[X]}(s)$ behave asymptotically as $T(s)^k$. Thus we are in a situation quite identical to the analysis of the profile of random DSTs in [3] and similar to the analysis of the profile of random tries in [17]. By the inverse Mellin transform [5], at $z = n$,

$$\tilde{M}_k^{[X]}(n) = \frac{1}{2\pi i} \int_{\rho-i\infty}^{\rho+i\infty} M_k^{*[X]}(s) n^{-s} \mathrm{d}s, \tag{21}$$

$$\tilde{M}_k^{[Y]}(n) = \frac{1}{2\pi i} \int_{\rho-i\infty}^{\rho+i\infty} M_k^{*[Y]}(s) n^{-s} \mathrm{d}s, \tag{22}$$

Here, we evaluate the integrals (21) and (22) via the saddle point method (see [6]). Thus, it is natural to choose $\rho = \rho_{n,k}$ as the saddle point of the function

$$T(s)^k n^{-s} = e^{k \log T(s) - s \log n},$$

i.e., the solution of $\frac{\partial}{\partial s}(k \log T(s) - s \log n) = 0$. Equivalently we must find ρ from

$$\alpha := \frac{k}{\log n} = \frac{p^{-\rho} + q^{-\rho}}{p^{-\rho}\log(1/p) + q^{-\rho}\log(1/q)},$$

that is, the only real-valued saddle point $\rho = \rho_{n,k} = \rho(\frac{k}{\log n})$,

$$\rho = \rho(\alpha) = \frac{1}{\log(p/q)} \log \frac{1-\alpha\log(1/p)}{\alpha\log(1/q)-1}, \qquad \big(\log(1/q)\big)^{-1} < \alpha < \big(\log(1/p)\big)^{-1}. \quad (23)$$

The integrands in (21) and (22), also has infinitely many complex-valued saddle points of the form $s_j := \rho + 2\pi i j/(\log p/q)$ $(j = \pm 1, \pm 2, \ldots)$. This is due to the fact

$$T(\rho + it) = p^{-\rho - it}\left(1 + \big(\frac{q}{p}\big)^{-\rho - it}\right) = p^{-\rho} \cdot e^{-it\log p}\left(1 + \big(\frac{q}{p}\big)^{-\rho} \cdot e^{it\log\frac{p}{q}}\right).$$

Now by putting $t = 2\pi j/(\log p/q)$, we have

$$\begin{aligned} T\big(\rho + 2\pi i j/(\log p/q)\big) &= p^{-\rho} \cdot e^{-2\pi i j(\log p)/(\log p/q)}\left(1 + \big(\frac{q}{p}\big)^{-\rho} \cdot e^{2\pi i j}\right) \\ &= e^{-2\pi i j(\log p)/(\log p/q)} T(\rho). \end{aligned}$$

Therefore the behaviour of $T(s)^k z^{-s}$ around $s = s_j$ is almost the same as that of $T(s)^k z^{-s}$ around $s = \rho$. This phenomenon gives a periodic leading factor in the asymptotics of $\tilde{M}_k^{[X]}(n)$ and $\tilde{M}_k^{[Y]}(n)$; and also of $\mathbb{E}(X_{n,k}) \approx \tilde{M}_k^{[X]}(n)$ and $\mathbb{E}(Y_{n,k}) \approx \tilde{M}_k^{[Y]}(n)$ (these two approximations are obtained by the analytical depoissonization given in [3], [6] and [8]).

If $\big(\log(1/p)\big)^{-1} + \varepsilon < \alpha < 2\big(\log(1/p) + \log(1/q)\big)^{-1} - \varepsilon$ (for some $\varepsilon > 0$); i.e., $\rho = \rho(\alpha) > 0$, then, by shifting the line of integration in (21) to the saddle point $\rho = \rho(\alpha)$, we collect a contribution of $F_k^{[X]}(0) = 2^k$ from the polar singularity of $\Gamma(s)F_k^{[X]}(s)$. This leads to

$$\tilde{M}_k^{[X]}(x) = 2^k + \frac{1}{2\pi i}\int_{\rho - i\infty}^{\rho + i\infty} M_k^{*[X]}(s) x^{-s} \mathrm{d}s. \quad (24)$$

Now, our main asymptotic results are given as follows.

THEOREM 3.5. Let $\mathbb{E}(X_{n,k})$ and $\mathbb{E}(Y_{n,k})$ denote the expectations of protected node profile and non-protected node profile in digital search trees with $0 < q := 1 - p < p < 1$.

Let $f^{[X]}(s)$ and $f^{[X]}(s)$ be the functions that are given by (16). Furthermore, we set

$$H^{[X]}(\rho,x) := \sum_{j\in\mathbb{Z}} \Gamma(\rho+it_j) f^{[X]}(\rho+it_j) e^{-2\pi ijx}, \qquad t_j := 2\pi j/(\log p/q),$$

$$H^{[Y]}(\rho,x) := \sum_{j\in\mathbb{Z}} \Gamma(\rho+it_j) f^{[Y]}(\rho+it_j) e^{-2\pi ijx}, \quad \beta(\rho) := \frac{p^{-\rho} q^{-\rho} \log(p/q)^2}{(p^{-\rho}+q^{-\rho})^2},$$

where $H^{[X]}(\rho,x)$ and $H^{[Y]}(\rho,x)$ are 1-periodic functions. Let k and n be positive integers and $\rho_{n,k} := \rho(\frac{k}{\log n})$ where $\rho(\alpha)$ is defined in (23), then we have:

(1) If $\frac{1}{\log\frac{1}{p}} + \varepsilon < \frac{k}{\log n} < \frac{2}{\log(1/p)+\log(1/q)} - \varepsilon$ (for some $\varepsilon > 0$), then uniformly

$$\mathbb{E}(X_{n,k}) = 2^k - H^{[X]}\big(\rho_{n,k}, \log_{p/q} p^k n\big) \frac{(p^{-\rho_{n,k}} + q^{-\rho_{n,k}})^k n^{-\rho_{n,k}}}{\sqrt{2\pi\beta(\rho_{n,k})k}} \big(1+\mathscr{O}\big(k^{-1/2}\big)\big)$$

(2) If $\frac{2}{\log(1/p)+\log(1/q)} + \varepsilon < \frac{k}{\log n} < \frac{1}{\log\frac{1}{q}} - \varepsilon$ (for some $\varepsilon > 0$), then uniformly

$$\mathbb{E}(X_{n,k}) = H^{[X]}\big(\rho_{n,k}, \log_{p/q} p^k n\big) \frac{(p^{-\rho_{n,k}} + q^{-\rho_{n,k}})^k n^{-\rho_{n,k}}}{\sqrt{2\pi\beta(\rho_{n,k})k}} \big(1+\mathscr{O}\big(k^{-1/2}\big)\big)$$

(3) If $\frac{1}{\log\frac{1}{p}} + \varepsilon < \frac{k}{\log n} < \frac{1}{\log\frac{1}{q}} - \varepsilon$ (for some $\varepsilon > 0$), then uniformly

$$\mathbb{E}(Y_{n,k}) = H^{[Y]}\big(\rho_{n,k}, \log_{p/q} p^k n\big) \frac{(p^{-\rho_{n,k}} + q^{-\rho_{n,k}})^k n^{-\rho_{n,k}}}{\sqrt{2\pi\beta(\rho_{n,k})k}} \big(1+\mathscr{O}\big(k^{-1/2}\big)\big)$$

PROOF. By evaluating the integrals (22) and (24) via the saddle point method; and using (8) and Lemma 3.4, the proof is quite identical to that of Theorem 4 and Theorem 5 in [3] with the new functions $f^{[X]}(s)$, $f^{[Y]}(s)$.

ACKNOWLEDGMENTS

We would like to thank Professor Michael Drmota for the comments on this research.

REFERENCES

L. Devroye and S. Janson, Protected nodes and fringe subtrees in some random trees, *Electron. Commun. Probab*, 19 (2014), 1–10.

M. Drmota, *Random trees: An interplay between combinatorics and probability*, SpringerWienNewYork, Vienna, 2009.

M. Drmota and W. Szpankowski, The expected profile of digital search trees, *J. Combin. Theory Ser. A*, 118 (2011), 1939–1965.

M. Drmota, M. Fuchs, H.K. Hwang and R. Neininger, Node profiles of symmetric digital search trees, https://arxiv.org/abs/1711.06941, submitted (2020), 41 pages.

P. Flajolet, X. Gourdon and P. Dumas, Mellin transforms and asymptotics: harmonic sums, *Theoret. Comput. Sci.*, 144 (1995), 3–58,.

P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, 2009.

M. Fuchs, C.-K. Lee and G.-R. Yu, On 2-protected nodes in random digital trees, *Theor. Comput. Sci.*, 622 (2016), 111–122.

P. Jacquet and W. Szpankowski, Analytical depoissonization and its applications, *Theoret. Comput. Sci.* 201 (1998), 1–62.

M. Javanian, Protected node profile of tries, *Discrete Mathematics & Theoretical Computer Science*, 20 (2018), 1–20.

M. Javanian, R. Imany Nabiyyi and J. Toofanpour, Protected node profile in recursive trees, *Theory of Probability and Its Applications*, accepted, 2021.

R. Kazemi and M. Vahidi-Asl, The variance of the profile in digital search trees, *Discrete Math. Theor. Comput. Sci.*, 13 (2011), 21–38.

D.E. Knuth, *The art of computer programming, Volume 3: sorting and searching*, Addison Wesley Longman Publishing Co., Inc. Redwood City, CA, USA, 1998.

Du. Rosena R.X. and H. Prodinger, Notes on protected nodes in digital search trees, *Appl. Math. Lett.*, 25 (2012), 1025–1028.

G. Louchard, Exact and asymptotic distributions in digital and binary search trees, *RAIRO Theor. Inform. Appl.*, 21 (1987), 479–495.

H. M. Mahmoud, *Evolution of random search trees*, John Wiley & Sons Inc., New York, 1992.

A. Magner, W. Szpankowski, Profile of PATRICIA tries, *Algorithmica*, 80 (2018), 331–397.

G. Park, H.-K. Hwang, P. Nicodème and W. Szpankowski, Profile of tries, *SIAM J. Comput.*, 38 (2009), 1821–1880.

W. Szpankowski, *Average case analysis of algorithms on sequences*, Wiley, New York, 2001.

V. M. Zolotarev, Approximation of the distributions of sums of independent random variables with values in infinite-dimensional spaces, *Theory Probab. Appl.*, 21 (1976), 721–737.

V. M. Zolotarev, Ideal metrics in the problem of approximating the distributions of sums of independent random variables, *Theory Probab. Appl.*, 22 (1977), 433–339,.

Mehri Javanian
University of Zanjan, Iran

Ramin Imany Nabiyyi
University of Tabriz, Iran

Jaleh Toofanpour
University of Tabriz, Iran

Mohammad Q. Vahidi-Asl
Shahid Beheshti University, Iran

Corresponding Author's address:
Mehri Javanian
Department of Statistics
University of Zanjan, Iran
Email: javanian@znu.ac.ir

Model choice for regression models with a categorical response

J. KALINA

Abstract

The multinomial logit model and the cumulative logit model represent two important tools for regression modeling with a categorical response with numerous applications in various fields. First, this paper presents a systematic review of these two models including available tools for model choice (model selection). Then, numerical experiments are presented for two real datasets with an ordinal categorical response. These experiments reveal that a backward model choice procedure by means of hypothesis testing is more effective compared to a procedure based on Akaike information criterion. While the tendency of the backward selection to be superior to Akaike information criterion has recently been justified in linear regression, such a result seems not to have been presented for models with a categorical response. In addition, we report a mistake in VGAM package of R software, which has however no influence on the process of model choice.

Mathematics Subject Classification 2010: 62J12, 62J02
General Terms: Regression, Generalized linear model, Model choice, Model selection, Information theory **Keywords:** categorical distribution, categorical response, Akaike information criterion.

1. INTRODUCTION

Regression models with a categorical response represent important tools for regression modeling with numerous applications in various fields including finance, econometrics, biomedicine, engineering, and others. We are interested in the generalized linear models (GLMs) that extend the logistic regression to situations with a non-binary response. Two (perhaps the most) important models are the multinomial logit model and the cumulative logit model; the former is appropriate for data with a nominal response and the latter for data with an ordinal response. These two models, which allow continuous and/or categorical regressors, have been described in the literature; still, the standard monographs on GLMs [Hosmer 2000; Agresti 2002] do not contain all useful details, which would be appreciated by users of statistical methods.

This work is supported by the project 21-19311S ("Information flow and equilibrium in financial markets") of the Czech Science Foundation.

10.2478/jamsi-2022-0005

Let us recall some remarkable recent applications of regression modeling with a categorical (non-binary) response. Housing demand was modeled in [Głuszak 2015] as a nominal 4-valued response, the values (levels) of which correspond to owning the property, renting a housing unit at market rate, renting it at reduced rate, or free accommodation. Travel and tourism in Aargau Canton in Switzerland was modeled as a nominal 4-valued response in [Wang et al. 2017], which corresponds to the 4 possibilities of traveling by car, transit, bicycle, or walking. Exit status of U.S. companies in the period 1976–1995 was modeled in [Irfan et al. 2018] as a nominal 4-valued response, the values of which correspond to voluntary liquidation of the company, involuntary liquidation, acquisition, or non-exit. Air quality index was modeled and forecasted as an ordinal response with 5 levels in [Kim 2017]. Finding the appropriate model, i.e. constructing a predictive model based only on the relevant predictors, represents an important task of regression modeling in all these applications.

Because this work is motivated by financial mathematics applications related to model selection for financial markets, let us pay particular attention also to financial applications of regression models with a categorical response. A cumulative logistic model was used in [Poplaski et al. 2019] to predict self-reported health outcomes of U.S. university students by their financial stress. A cross-nested logistic model for modeling an equilibrium within the cumulative prospect theory was formulated in [Yan and Yang 2021]. Financial distress of individual sectors of the U.S. economy was modeled by logistic regression as the response of macroeconomic variables in [Inekwe et al. 2018]. A logistic model was used to assess business financial health of Slovak companies in [Horváthová and Mokrišová 2020]. A functional logistic model was applied to predict anomalies of prices in the Chinese stock market in [Su et al. 2022].

We are interested in the multinomial logit model and cumulative logit model, which are the two most commonly used models for regression with a categorical response. In R software [R Core Team 2017], we are aware only of a single implementation of models with an ordinal categorical response. On the other hand, there are several recently created implementations of the multinomial logit model for a nominal response in R software. These are present in packages that contain other more advanced models tailor-made for specific situations. It turns out that there are several packages overlapping each other and there is a lack of comparisons of their test statistics. A numerical comparison of results obtained by different packages

Table I. Software packages in R available for fitting regression models with a categorical (non-binary) response, where the multinomial logit model (MLM) is a model for a nominal response and the cumulative logit model (CLM) is a model for an ordinal response.

Package	Response	Model	Function	Comments
VGAM	Nominal	MLM	vglm	Generalization of GLMs for a vector response
VGAM	Ordinal	CLM	cumulative	The package includes also other more advanced models
nnet	Nominal	MLM	multinom	Estimation by means of neural networks
mlogit	Nominal	MLM	mlogit	The package includes also other more advanced models
mnlogit	Nominal	MLM	mnlogit	Fast estimation of [Hasan et al. 2016]

seems missing and there seems no tutorial systematically describing and comparing the features of these packages. We present here Table I with the list of available R packages for a nominal or ordinal response.

This paper starts with a unique overview of the multinomial logit model (Section 2) and cumulative logit model (Section 3). Useful quantities including log-likelihoods or likelihood ratio test statistics are expressed in these methodological sections. Model choices for the two models are described in Section 4; these are not novel results, but remain difficult to be found in the literature. These include the popular backward selection based on hypothesis testing and Akaike information criterion, which appears less frequently in this context. Further, we present numerical experiments for two real datasets with an ordinal categorical response in Section 5; these are aimed at comparing backward selection by means of hypothesis testing and Akaike information criterion. The conclusions are presented in Section 6.

2. MULTINOMIAL LOGIT MODEL

We consider a nominal categorical variable $Y = (Y_1,\dots,Y_n)^T$, which is observed for the total number of n observations. Let us assume that the value of this response is equal to one of J values denoted as $v_1,\dots,v_J$. As Y is nominal (i.e. not ordinal), these values cannot be meaningfully ordered. Our task is to explain Y as a response of p-dimensional regressors (explanatory variables) $X_1,\dots,X_n$, which may be continuous or categorical (or a mixture of both). For the i-th observation, the regressor is denoted as $X_i = (X_{i1},\dots,X_{ip})^T \in \mathbb{R}^p$ for $i = 1,\dots,n$.

The multinomial logit model (MLM) [Agresti 2002] also known as the baseline-category logit model takes one category as the reference one. It has the form of a set of $J-1$ equations. If the J-th category is taken as the reference one, the model has the form

$$\log \frac{\mathsf{P}(Y_i = v_j)}{\mathsf{P}(Y_i = v_J)} = \alpha_j + \beta_{j1} X_{i1} + \cdots + \beta_{jp} X_{ip}, \quad i = 1, \ldots, n, \quad j = 1, \ldots, J-1. \quad (1)$$

This model for $J-1$ individual logits depends on the regression parameters

$$\alpha = (\alpha_1, \ldots, \alpha_{J-1})^T \in \mathbb{R}^{J-1}, \quad \beta_j = (\beta_{j1}, \ldots, \beta_{jp})^T \in \mathbb{R}^p \quad j = 1, \ldots, J-1. \quad (2)$$

From (1), we can easily obtain that

$$\mathsf{P}(Y_i = v_j) = \frac{\exp(\alpha_j + \beta_j^T X_i)}{1 + \sum_{h=1}^{J-1} \exp(\alpha_h + \beta_h^T X_i)}, \quad i = 1, \ldots, n, \quad j = 1, \ldots, J-1. \quad (3)$$

We now express the log-likelihood in the form

$$\begin{aligned} &\log L(\alpha, \beta_1, \ldots, \beta_{J-1}) = \log L(\alpha, \beta) = \\ &= \sum_{j=1}^{J-1} \left[\alpha_j \left(\sum_{i=1}^{n} \mathbb{1}[Y_i = v_j] \right) + \sum_{k=1}^{p} \beta_{jk} \left(\sum_{i=1}^{n} X_{ik} \mathbb{1}[Y_i = v_j] \right) \right] - \\ &- \sum_{i=1}^{n} \log \left[1 + \sum_{j=1}^{J-1} \exp(\alpha_j + \beta_j^T X_i) \right]. \end{aligned} \quad (4)$$

The parameters of the model can be estimated e.g. by the maximum likelihood method. It is convenient to obtain these estimates by maximizing the log-likelihood instead. Gradient-based optimization procedures, including the approach of Newton-Raphson, are fast and reliable in this task thanks to the concavity of (4) [Hasan et al. 2016]. A version of the multinomial logit model that allows to incorporate heterogeneity was proposed in [Tutz 2021].

3. CUMULATIVE LOGIT MODEL FOR AN ORDINAL RESPONSE

As in Section 2, we consider values of a categorical response $Y_1, \ldots, Y_n$ and values of p-dimensional regressors $X_1, \ldots, X_n$. This time, we assume Y to be ordinal. This means that Y attains one of J possible values denoted as $v_1, \ldots, v_J$, whose ordering is meaningful, i.e. we can use the fact that

$$v_1 \leq v_2 \leq \cdots \leq v_J. \quad (5)$$

It is therefore also meaningful to consider

$$P(Y_i \leq v_1) \leq \cdots \leq P(Y_i \leq v_J) = 1, \quad i = 1, \ldots, n. \quad (6)$$

Modeling such data with models assuming a continuous response was empirically investigated in [Fernández et al. 2020] and was strongly discouraged.

The most common regression model for this situation is the cumulative logit model (CLM) [Agresti 2002] described by a set of $J-1$ equations in the form

$$\log \frac{P(Y_i \leq v_j)}{1 - P(Y_i \leq v_j)} = \alpha_j + \beta^T X_i, \quad i = 1, \ldots, n, \quad j = 1, \ldots, J-1. \tag{7}$$

The model depends on parameters denoted as $\alpha = (\alpha_1, \ldots, \alpha_{J-1})^T$ and $\beta = (\beta_1, \ldots, \beta_p)^T$, where β is common for all $J-1$ equations. We can express the logarithmic likelihood as

$$\begin{aligned} \log L(\alpha, \beta) = \sum_{i=1}^{n} \Bigg\{ & \mathbb{1}[Y_i = v_1] \log \frac{\exp(\alpha_1 + \beta^T X_i)}{1 + \exp(\alpha_1 + \beta^T X_i)} + \\ & + \mathbb{1}[Y_i = v_J] \log \left(1 - \frac{\exp(\alpha_{J-1} + \beta^T X_i)}{1 + \exp(\alpha_{J-1} + \beta^T X_i)} \right) + \\ & + \sum_{j=2}^{J-1} \mathbb{1}[Y_i = v_j] \log \left(\frac{\exp(\alpha_j + \beta^T X_i)}{1 + \exp(\alpha_j + \beta^T X_i)} - \frac{\exp(\alpha_{j-1} + \beta^T X_i)}{1 + \exp(\alpha_{j-1} + \beta^T X_i)} \right) \Bigg\}. \end{aligned} \tag{8}$$

Bayesian estimation in cumulative logit models was recently studied in [Xu et al. 2022].

4. MODEL CHOICE

Numerous model choice approaches suitable for a continuous response including recent methods of [Ahrens et al. 2020; Shirk et al. 2018] cannot be easily extended to the model with a categorical response. The following methods seem to represent the major available approaches to model choice for models with a categorical response:

(A) Backward variable selection based on hypothesis testing, well known also from linear regression models [Kalina et al. 2019];

(B) Selecting the submodel with the minimal value of Akaike information criterion (AIC) over all subsets;

(C) Backward variable selection by means of AIC;

(D) Non-automatic backward variable selection based on hypothesis testing, i.e. including manual steps of the user, taking into account the interpretation of the final model.

(E) Regularized methods, i.e. regularized regression modeling for categorical response.

For the approach (A), any of asymptotic tests based on the likelihood function may be used, while the likelihood ratio test statistic (which will be expressed below) has actually the simplest form among them. Also none of the software packages overviewed in Table I reports any other tests besides the likelihood ratio test, which is denoted as residual deviance test there. Actually, the Wald test and Rao's score test [Rao 1973], which cannot be expressed analytically, can be hardly expressed in any elegant way because of their dependence of nuisance parameters.

Let us consider testing H_0, which states that there is no effect of a given regressor (or their block) on the response, against H_1, which states the contrary of H_0. The likelihood ratio test statistic G^2, commonly denoted as deviance test (residual deviance test) in the context of GLMs, can be expressed as

$$LR = G^2(H_0|H_1) = 2\left[\log L(H_1) - \log L(H_0)\right] = -2\left[\log L(H_0) - \log L(H_1)\right]; \qquad (9)$$

here, L denotes the likelihood function evaluated for all the observations and the remaining notation is already self-explaining. Under H_0, (9) has an asymptotic χ^2 distribution, where the number of degrees of freedom corresponds to the difference between the numbers of parameters under H_1 and H_0. The test rejects H_0 if (9) is too large, i.e. exceeds the corresponding critical value. The LR test statistic is a measure of goodness of fit of the submodel with the broader model. We can say that the LR test compares the maximum possible value of L (or $\log L$) under the submodel with its maximum possible value attainable under the broader model.

The backward selection process based on the LR testing builds the model sequentially, starting with a general model and proceeding to a (possibly) simple submodel with an insignificant value of the likelihood ratio test statistic. In each particular step of the model building, the null hypothesis corresponds to the possibility to proceed from a model that has already been approved as suitable to its particular submodel.

As an alternative to testing, it is possible to consider Akaike information criterion (AIC), which represents a general information-theoretical measure of quality of a regression fit, tailor-made for model selection [Akaike 1973]. For a given multinomial logit model or cumulative logit model, AIC can be schematically expressed as

$$\mathsf{AIC} = 2k - 2\log L(\hat{\alpha}, \hat{\beta}), \qquad (10)$$

where k is the number of parameters in the model and L represents the likelihood evaluated for maximum likelihood estimates. Model choice can be performed by selecting the model with the minimal value of AIC over all possible (suitable) models. It is surprising that AIC as a tool for model choice is not at all mentioned in standard monographs devoted to regression models with a categorical response [Hosmer 2000; Agresti 2002]. The approach (B) is common in linear regression, while the idea of (C) is to simplify the computations especially for a large p [Harrell 2015].

4.1. Model choice for the multinomial logit model

In the model of Section 2, let us consider the null hypothesis $H_0 : \beta_{1m} = \cdots = \beta_{J-1,m} = 0$ for a particular $m = 1,\ldots,p$. This corresponds to testing that there is no effect of the m-th regressor on the response, conditionally on the effect of the remaining regressors. The maximum likelihood estimates of α and β will be denoted here as $\hat{\alpha}$ and $\hat{\beta} = (\hat{\beta}_1,\ldots,\hat{\beta}_{J-1})^T$. The logarithmic likelihood under H_1 is obtained by plugging $\hat{\alpha}$ and $\hat{\beta}$ into (4), i.e. equals $\log L(\hat{\alpha},\hat{\beta})$. Let us further use the notation $\tilde{\alpha}_j$ and $\tilde{\beta}_k$ for maximum likelihood estimates obtained under H_0, i.e. obtained so that $\tilde{\beta}_{1m} = \cdots = \tilde{\beta}_{J-1,m} = 0$. The logarithmic likelihood ratio under H_0 is obtained as

$$\begin{aligned} \log L(\alpha^0,\beta^0) = \sum_{j=1}^{J-1} \left[\tilde{\alpha}_j \left(\sum_{i=1}^{n} Y_{ij} \right) + \sum_{m\neq k=1}^{p} \tilde{\beta}_{jk} \left(\sum_{i=1}^{n} X_{ik} \mathbb{1}[Y_i = v_j] \right) \right] - \\ - \sum_{i=1}^{n} \log \left[1 + \sum_{j=1}^{J-1} \exp \left(\tilde{\alpha}_j + \sum_{m\neq k=1}^{p} \tilde{\beta}_{jk} X_{ik} \right) \right]. \end{aligned} \tag{11}$$

The likelihood ratio test statistic can be in our notation expressed as

$$LR = -2 \left[\log L(\alpha^0,\beta^0) - \log L(\hat{\alpha},\hat{\beta}) \right]. \tag{12}$$

Under H_0, the test statistic has asymptotically the χ^2 distribution with $J-1$ degrees of freedom.

Because the multinomial logit model contains $p(J-1)$ parameters, AIC for the model can be expressed as

$$\mathsf{AIC} = 2p(J-1) - 2\log L(\hat{\alpha},\hat{\beta}). \tag{13}$$

4.2. Model choice for the cumulative logit model

In the model of Section 3, let us consider the null hypothesis $H_0 : \beta_m = 0$ for a fixed $m \in \{1,\ldots,p\}$. The maximum likelihood estimates of the parameters will be denoted as $\hat{\alpha}$ and $\hat{\beta}$. The logarithmic likelihood under H_1, which is equal to

$\log L(\hat{\alpha},\hat{\beta})$ expressed in (8), is obtained as the log-likelihood that uses maximum likelihood estimates of the parameters. The logarithmic likelihood under H_0 can be expressed as $\log L(\alpha^0,\beta^0)$, where $\tilde{\alpha}_j$ and $\tilde{\beta}_k$ are maximum likelihood estimates obtained under H_0. Particularly, it holds $\tilde{\beta}_m = 0$. Again, we can construct the likelihood ratio test statistic as in (12). Such statistic has asymptotically the χ^2 distribution with $J-1$ degrees of freedom under H_0.

Because the cumulative logit model contains $p+J-1$ parameters, AIC for the model can be expressed as

$$\mathsf{AIC} = 2\,(p+J-1) - 2\log L(\hat{\alpha},\hat{\beta}). \tag{14}$$

5. NUMERICAL EXPERIMENTS

To illustrate the model choice approaches of Section 4, we consider two datasets obtained as two disjoint parts of the Wine Quality Dataset, which was originally presented in [Cortez et al. 2009] and is now publicly available [Dua and Graff 2017]. We consider the part of the data corresponding to white wine as one dataset, and the part corresponding to red wine as the other dataset. Both datasets consider the wine quality to be the response, which is an ordinal variable with 7 possible values, namely integers in the set $\{3,4,5,6,7,8,9\}$. There are 11 regressors in both datasets, where the white wine dataset contains $n = 4898$ observations and the red wine dataset contains $n = 1599$ observations. We use the package VGAM [Yee 2010] of R software for all computations; particularly, the following code works for modeling an ordinal response, where x denotes the matrix of regressors.

```
library(VGAM);
fit = vglm(y~x, family=cumulative(parallel=TRUE))
summary(fit);
```

First, we consider the white wine dataset. The results as well as the methods used for the computation are presented in Table II. The backward selection (A) reduces the model with 11 regressors to the model with 9 relevant regressors presented in the table, while it is not possible to reduce it to any model with 8 regressors as the likelihood ratio test is always significant. We do not report point estimates of the parameters of individual models, because it is commonly more important in applications to interpret which variables are those contributing to the variability of the response. If AIC is used in (B) and (C), the resulting model is the same as under (A).

Further, we analyze the red wine dataset and present the results in Table III.

The backward selection (A) reduces the model with 11 regressors to the model with 7 relevant regressors presented in the table. Here, it is not possible to reduce such model to any model with 6 regressors, because the likelihood ratio test statistic would be significant in every such situation. Approaches (B) and (C) yield a model with 9 regressors, which is one of submodels (but not the final one) considered during the computation of (A). The computation of (B) requires to consider $2^p - 1$ submodels of the model with all regressors, which means that the log-likelihood is evaluated in 2047 models for each of our datasets. The complexity of the computation of (A) heavily depends on a particular dataset. The computation of (A) required here to evaluate the log-likelihood twice for each model (i.e. under H_0 and H_1) and was thus much smaller here compared to the complexity of (B).

Let us recall a recent result of [Heinze et al. 2018] revealing AIC to be inferior to backward selection by means of testing (A) in linear regression for data with a large n/p ratio. Particularly, [Heinze et al. 2018] recommended (A) in linear regression if the number of events-per-variable (EPV) exceeds 25. Here, the white wine dataset has $\mathsf{EPV} = 4898/11 \doteq 445$ and the red wine dataset has $\mathsf{EPV} = 1599/11 \doteq 145$. In other words, the datasets are very large in terms of EPV. This may explain (B) and (C) to be weak, although such performance seems not to have been reported before for models with a categorical response.

We found a mistake in the vglm function of VGAM package. It turns out in Tables II and III that residual deviance is reported to be precisely equal to the (-2)-multiple of the reported logarithmic likelihood. This cannot be however true, which is clear from the definition of residual deviance. We verified that this mistake does not appear for the simpler situation in the glm function of stats package of R software, but there is no documentation explaining this for the VGAM package. In the analysis of this phenomenon, we verified the residual deviance to be correct and log-likelihood incorrect in VGAM. Still, we recommend the users to compute AIC directly from the definition (14) instead of relying on the log-likelihood presented in the VGAM package.

6. CONCLUSION

The paper is devoted to the multinomial logit model and cumulative logit model. The methodological part recalls the models and expresses useful quantities such as log-likelihoods, likelihood ratio test statistics, or Akaike information criterion. These can be hardly found in the available literature. Therefore, this paper may potentially

Table II. Results of the analysis for the white wine dataset. For each model, the number of parameters, log-likelihood, residual deviance, and AIC are reported.

Model with regressors Source of computation	# of par. Own	Log-lik. [Yee 2010]	Residual deviance [Yee 2010]	AIC Own (14)
(A) Backward selection based on the LR test				
1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11	17	−5450.4	10 900.9	10 934.8
1, 2, 3, 4, 6, 7, 8, 9, 10, 11	16	−5450.5	10 901.1	10 933.0
1, 2, 4, 6, 7, 8, 9, 10, 11	15	−5450.7	10 901.3	10 931.4
(B) Minimal AIC over all subsets				
1, 2, 4, 6, 7, 8, 9, 10, 11	15	−5450.7	10 901.3	10 931.4
(C) Backward selection based on AIC				
1, 2, 4, 6, 7, 8, 9, 10, 11	15	−5450.7	10 901.3	10 931.4

Table III. Results of the analysis for the red wine dataset.

Model with regressors Source of computation	# of par. Own	Log-lik. [Yee 2010]	Residual deviance [Yee 2010]	AIC Own (14)
(A) Backward selection based on the LR test				
1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11	17	−1537.4	3 074.8	3 108.8
1, 2, 3, 4, 5, 6, 7, 9, 10, 11	16	−1538.0	3 076.0	3 108.0
2, 3, 4, 5, 6, 7, 9, 10, 11	15	−1538.3	3 076.6	3 106.6
2, 4, 5, 6, 7, 9, 10, 11	14	−1539.5	3 079.1	3 107.0
2, 5, 6, 7, 9, 10, 11	13	−1540.4	3 080.8	3 108.8
(B) Minimal AIC over all subsets				
2, 3, 4, 5, 6, 7, 9, 10, 11	15	−1538.3	3 076.6	3 106.6
(C) Backward selection based on AIC				
2, 3, 4, 5, 6, 7, 9, 10, 11	15	−1538.3	3 076.6	3 106.6

find applications as a learning material for an advanced course on GLMs. In case that $J = 2$, all formulas presented in the whole paper are precisely equal to those of the logistic regression.

The main contribution of the paper is the numerical experiments, which illustrate different model choice approaches for the two regression models with a categorical response. During the computations, we found a mistake in the VGAM package of R software, which does not however influence any of the model choice approaches. AIC turns out to be less efficient compared to the backward stepwise procedure, as it keeps also statistically insignificant variables in the final model. The backward selection based on testing is of course at the same time computationally more appealing compared to computing AIC over all possible subsets of regressors.

It has been realized only recently that AIC remains less effective compared to backward selection based on hypothesis testing in linear regression [Heinze et al. 2018]; this result seems to be little known among practitioners and definitely going much beyond the experience described in the fundamental monograph [Harrell 2015]. Our results evaluated over two datasets show that an analogous result may be true also for regression with a categorical response, which would deserve to be examined over a much larger number of diverse datasets. In any case, it is very practical to know that the result of the hierarchical model choice procedure is stable; instability has namely been known as the major drawback of many standard model choice procedures [Breiman 1996].

Other more advanced models, e.g. the mixed logit models of [Sarrias and Daziano 2017] also known as random coefficients multinomial logit models, exceed the scope of this paper. The authors intend to investigate model choice procedures and to focus on their Bayesian versions in the context of financial modeling. To mention some open problems for models with a categorical response, we are not aware of regression quantiles or minimum redundance maximum relevance (MRMR) variable selection approaches [Kalina and Schlenker 2015].

ACKNOWLEDGEMENTS

The author would like to thank an anonymous referee for valuable suggestions and Anežka Faltýnková for technical help.

REFERENCES

AGRESTI, A. 2002. *Categorical data analysis*, 2 ed. Wiley, Hoboken.

AHRENS, A., HANSEN, C., AND SCHAFFER, M. 2020. Lassopack: Model selection and prediction with regularized regression in §tata. *The Stata Journal 20*, 176–235.

AKAIKE, H. 1973. A new look at the statistical model identification. *IEEE Transactions of Automatic Control 19*, 716–723.

BREIMAN, L. 1996. Heuristics of instability and stabilization in model selection. *Annals of Statistics 24*, 2350–2383.

CORTEZ, P., CERDEIRA, A., ALMEIDA, F., MATOS, T., AND REIS, J. 2009. Modeling wine preferences by data mining from physicochemical properties. *Decision Support Systems 47*, 547–553.

DUA, D. AND GRAFF, C. 2017. UCI machine learning repository.

FERNÁNDEZ, D., LIU, I., COSTILLA, R., AND GU, P. 2020. Assigning scores for ordered categorical responses. *Journal of Applied Statistics 47*, 1261–1281.

GŁUSZAK, M. 2015. Multinomial logit model of housing demand in Poland. *Real Estate Management and Valuation 23*, 84–89.

HARRELL, F. 2015. *Regression modeling strategies with applications to linear models, logistic and ordinal regression, and survival analysis*, 2 ed. Springer, Cham.

HASAN, A., WANG, Z., AND MAHANI, A. 2016. Fast estimation of multinomial logit models: R package mnlogit. *Journal of Statistical Software 75*, 1–24.

HEINZE, G., WALLISCH, C., AND DUNKLER, D. 2018. Variable selection—a review and recommendations for the practicing statistician. *Biometrical Journal 60*, 431–449.

HORVÁTHOVÁ, J. AND MOKRIŠOVÁ, M. 2020. Comparison of the results of a data envelopment analysis model and logit model in assessing business financial health. *Information 11*, 160.

HOSMER, D. 2000. *Applied logistic regression*, 2 ed. Wiley, Hoboken.

INEKWE, J., JIN, Y., AND VALENZUELA, M. 2018. The effects of financial distress: Evidence from US GDP growth. *Economic Modelling 72*, 8–21.

IRFAN, M., SAHA, S., AND SINGH, S. 2018. A random effects multinomial logit model for the determinants of exit modes. evidence from a panel of us manufacturing firms. *Journal of Economic Studies 45*, 791–809.

KALINA, J. AND SCHLENKER, A. 2015. A robust supervised variable selection for noisy high-dimensional data. *BioMed Research International 2015*, 320385.

KALINA, J., VAŠANIČOVÁ, P., AND LITAVCOVÁ, E. 2019. Regression quantiles under heteroscedasticity and multicollinearity: Analysis of travel and tourism competitiveness. *Ekonomický časopis/Journal of Economics 67*, 69–85.

KIM, S. 2017. Ordinal time series model for forecasting air quality index for ozone in southern california. *Environmental Modeling & Assessment 22*, 175–182.

POPLASKI, S., KEMNITZ, R., AND ROBB, C. 2019. Investing in education: Impact of student financial stress on self-reported health. *Journal of Student Financial Aid 48*, 3.

R CORE TEAM. 2017. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

RAO, C. 1973. *Linear statistical inference and its applications*, 2 ed. Wiley, New York.

SARRIAS, M. AND DAZIANO, R. 2017. Multinomial logit models with continuous and discrete individual heterogeneity in r: The gmnl package. *Journal of Statistical Software 19*, 1–46.

SHIRK, A., LANDGUTH, E., AND CUSHMAN, S. 2018. A comparison of regression methods for model selection in individual-based landscape genetic analysis. *Molecular Ecology Resources 18*, 55–67.

SU, Z., BAO, H., LI, Q., XU, B., AND CUI, X. 2022. The prediction of price gap anomaly in Chinese stock market: Evidence from the dependent functional logit model. *Finance Research Letters*. In press.

TUTZ, G. 2021. Uncertain choices: The heterogeneous multinomial logit model. *Sociological Methodology 51*, 86–111.

WANG, K., YE, X., PENDYALA, R., AND ZOU, Y. 2017. On the development of a semi-nonparametric generalized multinomial logit model for travel-related choices. *PLoS ONE 12*, e0186689.

XU, J., BAULDRY, S., AND FULLERTON, A. 2022. Bayesian approaches to assessing the parallel lines assumption in cumulative ordered logit models. *Sociological Methods & Research*. In press.

YAN, D. AND YANG, Y. 2021. A stochastic user equilibrium formulation for the cumulative prospect theory-based cross-nested logit. *Discrete Dynamics in Nature and Society 2021*, 9929015.

YEE, T. W. 2010. The VGAM package for categorical data analysis. *Journal of Statistical Software 32,* 10, 1–34.

Jan Kalina
The Czech Academy of Sciences,
Institute of Computer Science, Prague
&
Charles University,
Faculty of Mathematics and Physics, Prague
Email: kalina@cs.cas.cz

Fractional Simpson like type inequalities for differentiable s-convex functions

N. KAMOUCHE, S. GHOMRANI AND B. MEFTAH

Abstract

In this paper, based on new identity we establish some new Simpson like type inequalities for functions whose first derivatives are s-convex via Riemann-Liouville fractional integrals. The case where the derivatives are bounded as well as the case where the derivatives satisfy the Hölder condition are also discussed. The obtained results extend some known results and refine another one. Applications of the results are given at the end.

Mathematics Subject Classification 2000: 26D10, 26D15, 26A51
Keywords: Simpson like inequality, s-convex functions, Hölderian functions, bounded functions.

1. INTRODUCTION

Let I be an interval of real numbers

Definition 1.1. [15] A function $f : I \to \mathbb{R}$ is said to be convex, if

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

holds for all $x, y \in I$ and all $t \in [0,1]$.

The theory of convexity plays a central and attractive role in many fields of research. This theory provides us with a powerful tool for solving a large class of problems that arise in pure and applied mathematics. In recent years, the concept of convexity has been improved, generalized and extended in many directions. Among these generalizations, we note that introduced by Breckner called s-convexity and can be defined as follows

Definition 1.2. [6] A nonnegative function $f : I \subset [0,\infty) \to \mathbb{R}$ is said to be s-convex in the second sense for some fixed $s \in (0,1]$, if

$$f(tx + (1-t)y) \leq t^s f(x) + (1-t)^s f(y)$$

holds for all $x, y \in I$ and $t \in [0,1]$.

10.2478/jamsi-2022-0006

The concept of convexity has a close relationship in the development of the theory of inequalities. The combination of these theories has attracted a lot of attention from researchers due to their nature and properties.

One of the most important and widely requested inequalities is that of Simpson which can be stated as follows

$$\left|\tfrac{1}{6}\left(f(a)+4f\left(\tfrac{a+b}{2}\right)+f(b)\right)-\tfrac{1}{b-a}\int_a^b f(u)\,du\right|\leq \tfrac{(b-a)^4}{2880}\left\|f^{(4)}\right\|_\infty, \tag{1}$$

where f is four-times continuously differentiable function on (a,b), and $\left\|f^{(4)}\right\|_\infty=\sup\limits_{x\in(a,b)}\left|f^{(4)}(x)\right|$.

Concerning some papers related to the inequality (1) see [1; 2; 3; 4; 5; 7; 8; 9; 10; 13; 16; 18], and references therein.

Recently, Shuang et al. [17] discussed an inequality similar to (1), Among the obtained results for convex functions, we quote

$$\begin{aligned}&\left|\tfrac{1}{8}\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right)-\tfrac{1}{b-a}\int_a^b f(t)\,dt\right|\\ \leq\ & \tfrac{b-a}{768}\left(19\left|f'(a)\right|+82\left|f'\left(\tfrac{a+b}{2}\right)\right|+19\left|f'(b)\right|\right),\end{aligned}$$

$$\begin{aligned}&\left|\tfrac{1}{8}\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right)-\tfrac{1}{b-a}\int_a^b f(t)\,dt\right|\\ \leq\ & \tfrac{b-a}{16}\left(\tfrac{1+3^{p+1}}{4(p+1)}\right)^{\frac{1}{p}}\left(\left(\tfrac{\left|f'(a)\right|^q+\left|f'\left(\frac{a+b}{2}\right)\right|^q}{2}\right)^{\frac{1}{q}}+\left(\tfrac{\left|f'\left(\frac{a+b}{2}\right)\right|^q+\left|f'(b)\right|^q}{2}\right)^{\frac{1}{q}}\right)\end{aligned}$$

and

$$\begin{aligned}&\left|\tfrac{1}{8}\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right)-\tfrac{1}{b-a}\int_a^b f(t)\,dt\right|\\ \leq\ & \tfrac{5(b-a)}{64}\left(\left(\tfrac{19\left|f'(a)\right|^q+41\left|f'\left(\frac{a+b}{2}\right)\right|^q}{60}\right)^{\frac{1}{q}}+\left(\tfrac{41\left|f'\left(\frac{a+b}{2}\right)\right|^q+19\left|f'(b)\right|^q}{60}\right)^{\frac{1}{q}}\right).\end{aligned}$$

In [14], Luo et al. gave the analogue weighted version of the result given by Shuang et al. [14]. They also discussed the cases where the first derivatives are bounded and Lipschitzian, of which we quote two of the established results: For functions whose first derivatives are bounded i.e. $m\leq f(x)\leq M$ for all $x\in[a,b]$, we have

$$\left|\tfrac{1}{8}\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right)-\tfrac{1}{b-a}\int_a^b f(t)\,dt\right|\leq\tfrac{9(b-a)(M-m)}{64}.$$

For functions whose first derivatives satisfies a Lipschitz condition i.e. $\left|f'(x)-f'(y)\right|\leq$

$L|x-y|$ for all $x,y\in[a,b]$, we have

$$\left|\frac{1}{8}\left(f(a)+6f\left(\frac{a+b}{2}\right)+f(b)\right)-\frac{1}{b-a}\int_a^b f(t)\,dt\right|\leq\frac{41(b-a)^2L}{768}+\frac{b-a}{16}\left(f'(a)+f'(b)\right).$$

Very recently, Kirmaci [12], gave the following inequalities for s-convex and convex first derivatives

$$\begin{aligned}&\left|\frac{1}{8}\left(f(a)+6f\left(\frac{a+b}{2}\right)+f(b)\right)-\frac{1}{b-a}\int_a^b f(t)\,dt\right|\\ \leq\ &\frac{(b-a)\left(3^{p+1}+1\right)^{\frac{1}{p}}}{8\times16^{\frac{1}{p}}}\left(\frac{1}{2^{\frac{s+1}{q}}}+\left(1-\frac{1}{2^{s+1}}\right)^{\frac{1}{q}}\right)\left(|f'(a)|+|f'(b)|\right)\end{aligned}$$

and

$$\begin{aligned}&\left|\frac{1}{8}\left(f(a)+6f\left(\frac{a+b}{2}\right)+f(b)\right)-\frac{1}{b-a}\int_a^b f(t)\,dt\right|\\ \leq\ &\frac{(b-a)\left(3^{p+1}+1\right)^{\frac{1}{p}}}{16\times8^{\frac{1}{p}}}\left(|f'(a)|+2\left|f'\left(\frac{a+b}{2}\right)\right|+|f'(b)|\right).\end{aligned}$$

In last decades the fractional calculus has attracted the attention of many researchers due to its has wide applications in pure and applied mathematics, especially the Riemann-Liouville operator which we recall the definition

Definition 1.3. [11] Let $f\in L^1[a,b]$. The Riemann-Liouville fractional integrals $I^{\alpha}_{a^+}f$ and $I^{\alpha}_{b^-}f$ of order $\alpha>0$ with $a\geq0$ are defined by

$$\begin{aligned}I^{\alpha}_{a^+}f(x)&=\frac{1}{\Gamma(\alpha)}\int_a^x(x-t)^{\alpha-1}f(t)dt,\quad x>a,\\ I^{\alpha}_{b^-}f(x)&=\frac{1}{\Gamma(\alpha)}\int_x^b(t-x)^{\alpha-1}f(t)dt,\quad b>x,\end{aligned}$$

respectively. Here $\Gamma(\alpha)=\int_0^{\infty}e^{-t}t^{\alpha-1}dt$, is the gamma function and $I^0_{a^+}f(x)=I^0_{b^-}f(x)=f(x)$.

Motivated by the results cited above, the aim of this study is to extend the results given in [17] for functions whose first derivatives are s-convex via Riemann-Liouville fractional integrals. We also discuss the cases where the derivatives are bounded and satisfy the Hölder condition. The results obtained are based on a new fractional identity and refine those of [12; 14]. We end the paper with a few applications.

2. MAIN RESULTS

In order to prove our results, we need the following lemma

LEMMA 2.1. Let $f : I \subset \mathbb{R} \to \mathbb{R}$ be a differentiable function on I°, $a, b \in I^\circ$ with $a < b$, and $f' \in L^1[a,b]$, then the following equality holds:

$$\begin{aligned}
&\tfrac{1}{8}\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right)-\tfrac{\Gamma(\alpha+1)}{2^{1-\alpha}(b-a)^\alpha}\left(I^\alpha_{\left(\frac{a+b}{2}\right)^-}f(a)+I^\alpha_{\left(\frac{a+b}{2}\right)^+}f(b)\right)\\
= &\tfrac{b-a}{4}\left(\int_0^1\left(t^\alpha-\tfrac{1}{4}\right)f'\left((1-t)a+t\tfrac{a+b}{2}\right)dt\right.\\
&\left.-\int_0^1\left((1-t)^\alpha-\tfrac{1}{4}\right)f'\left((1-t)\tfrac{a+b}{2}+tb\right)dt\right).
\end{aligned}$$

PROOF. Let

$$\begin{aligned}
I_1 &= \int_0^1\left(t^\alpha-\tfrac{1}{4}\right)f'\left((1-t)a+t\tfrac{a+b}{2}\right)dt\\
I_2 &= \int_0^1\left(\tfrac{1}{4}-(1-t)^\alpha\right)f'\left((1-t)\tfrac{a+b}{2}+tb\right)dt.
\end{aligned}$$

Integrating by parts I_1, we get

$$\begin{aligned}
I_1 = &\tfrac{2}{b-a}\left(t^\alpha-\tfrac{1}{4}\right)f\left((1-t)a+t\tfrac{a+b}{2}\right)\Big|_{t=0}^{t=1}\\
&-\tfrac{2\alpha}{b-a}\int_0^1 t^{\alpha-1}f\left((1-t)a+t\tfrac{a+b}{2}\right)dt\\
= &\tfrac{3}{2(b-a)}f\left(\tfrac{a+b}{2}\right)+\tfrac{1}{2(b-a)}f(a)-\tfrac{2\alpha}{b-a}\int_0^1 t^{\alpha-1}f\left((1-t)a+t\tfrac{a+b}{2}\right)dt\\
= &\tfrac{3}{2(b-a)}f\left(\tfrac{a+b}{2}\right)+\tfrac{1}{2(b-a)}f(a)-\alpha\left(\tfrac{2}{b-a}\right)^{\alpha+1}\int_a^{\frac{a+b}{2}}(u-a)^{\alpha-1}f(u)\,du\\
= &\tfrac{3}{2(b-a)}f\left(\tfrac{a+b}{2}\right)+\tfrac{1}{2(b-a)}f(a)-\tfrac{2^{\alpha+1}\Gamma(\alpha+1)}{(b-a)^{\alpha+1}}I^\alpha_{\left(\frac{a+b}{2}\right)^-}f(a).
\end{aligned}\tag{2}$$

Similarly, we get

$$\begin{aligned}
I_2 = &\tfrac{2}{b-a}\left((1-t)^\alpha-\tfrac{1}{4}\right)f\left((1-t)\tfrac{a+b}{2}+tb\right)\Big|_{t=0}^{t=1}\\
&+\tfrac{2\alpha}{b-a}\int_0^1(1-t)^{\alpha-1}f\left((1-t)\tfrac{a+b}{2}+tb\right)dt\\
= &-\tfrac{1}{2(b-a)}f(b)-\tfrac{3}{2(b-a)}f\left(\tfrac{a+b}{2}\right)+\tfrac{2\alpha}{b-a}\int_0^1(1-t)^{\alpha-1}f\left((1-t)\tfrac{a+b}{2}+tb\right)dt\\
= &-\tfrac{1}{2(b-a)}f(b)-\tfrac{3}{2(b-a)}f\left(\tfrac{a+b}{2}\right)+\alpha\left(\tfrac{2}{b-a}\right)^{\alpha+1}\int_{\frac{a+b}{2}}^{b}(b-u)^{\alpha-1}f(u)\,du\\
= &-\tfrac{1}{2(b-a)}f(b)-\tfrac{3}{2(b-a)}f\left(\tfrac{a+b}{2}\right)+\tfrac{2^{\alpha+1}\Gamma(\alpha+1)}{(b-a)^{\alpha+1}}I^\alpha_{\left(\frac{a+b}{2}\right)^+}f(b).
\end{aligned}\tag{3}$$

Making the difference between (2) and (3), and then multiplying the resulting equality by $\frac{b-a}{4}$, we get the desired result.

Before giving our main results, let's recall some special functions that we will call in the sequel

Definition 2.2. [11] For any complex numbers and nonpositive integers x,y such that $\operatorname{Re}(x)>0$ and $\operatorname{Re}(y)>0$. The beta function is defined by

$$B(x,y)=\int_0^1 t^{x-1}(1-t)^{y-1}\,dt=\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)},$$

where $\Gamma(.)$ is the Euler gamma function. The incomplete beta function.

Definition 2.3. [11] For any complex numbers and nonpositive integers x,y such that $\operatorname{Re}(x)>0$ and $\operatorname{Re}(y)>0$, we have

$$B_a(x,y)=\int_0^a t^{x-1}(1-t)^{y-1}\,dt,\ \ a<1.$$

Definition 2.4. [11] The hypergeometric function is defined for $\operatorname{Re}c>\operatorname{Re}b>0$ and $|z|<1$, as follows

$${}_2F_1(a,b,c;z)=\frac{1}{B(b,c-b)}\int_0^1 t^{b-1}(1-t)^{c-b-1}(1-zt)^{-a}\,dt,$$

where $c>b>0, |z|<1$ and $B(.,.)$ is the beta function.

THEOREM 2.5. Let $f:[a,b]\to\mathbb{R}$ be a differentiable function on (a,b) such that $f'\in L^1[a,b]$ with $0\le a<b$. If $|f'|$ is s-convex in the second sense for some fixed $s\in(0,1]$, then we have

$$\begin{aligned}&\left|\frac{1}{8}\left(f(a)+6f\left(\frac{a+b}{2}\right)+f(b)\right)-\frac{\Gamma(\alpha+1)}{2^{1-\alpha}(b-a)^\alpha}\left(I^\alpha_{\left(\frac{a+b}{2}\right)^-}f(a)+I^\alpha_{\left(\frac{a+b}{2}\right)^+}f(b)\right)\right|\\ \le\ &\frac{b-a}{4}\left(\Theta_{s,\alpha}|f'(a)|\right.\\ &\left.+\left(\frac{\alpha}{(s+1)(\alpha+s+1)}\left(\frac{1}{4}\right)^{\frac{s+1}{\alpha}}+\frac{3s+3-\alpha}{2(s+1)(\alpha+s+1)}\right)\left|f'\left(\frac{a+b}{2}\right)\right|+\Theta_{s,\alpha}|f'(b)|\right),\end{aligned}$$

where

$$\begin{aligned}\Theta_{s,\alpha}=\ &\frac{1}{4(s+1)}\left(1-2\left(1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}\right)^{s+1}\right)\\ &+B_{1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}(s+1,\alpha+1)-B_{\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}(\alpha+1,s+1).\end{aligned}\tag{4}$$

PROOF. From Lemma 2.1, properties of modulus, and s-convexity in the second sense of $|f'|$, we have

$$
\begin{aligned}
&\left|\tfrac{1}{8}\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right)-\tfrac{\Gamma(\alpha+1)}{2^{1-\alpha}(b-a)^{\alpha}}\left(I^{\alpha}_{\left(\frac{a+b}{2}\right)^{-}}f(a)+I^{\alpha}_{\left(\frac{a+b}{2}\right)^{+}}f(b)\right)\right|\\
\leq\ & \tfrac{b-a}{4}\left(\int_0^1\left|t^{\alpha}-\tfrac{1}{4}\right|\left|f'\left((1-t)a+t\tfrac{a+b}{2}\right)\right|dt\right.\\
&\left.+\int_0^1\left|(1-t)^{\alpha}-\tfrac{1}{4}\right|\left|f'\left((1-t)\tfrac{a+b}{2}+tb\right)\right|dt\right)\\
=\ & \tfrac{b-a}{4}\left(\int_0^{\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}\left(\tfrac{1}{4}-t^{\alpha}\right)\left|f'\left((1-t)a+t\tfrac{a+b}{2}\right)\right|dt\right.\\
&+\int_{\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}^{1}\left(t^{\alpha}-\tfrac{1}{4}\right)\left|f'\left((1-t)a+t\tfrac{a+b}{2}\right)\right|dt\\
&+\int_0^{1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}\left((1-t)^{\alpha}-\tfrac{1}{4}\right)\left|f'\left((1-t)\tfrac{a+b}{2}+tb\right)\right|dt\\
&\left.+\int_{1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}^{1}\left(\tfrac{1}{4}-(1-t)^{\alpha}\right)\left|f'\left((1-t)\tfrac{a+b}{2}+tb\right)\right|dt\right)\\
\leq\ & \tfrac{b-a}{4}\left(\int_0^{\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}\left(\tfrac{1}{4}-t^{\alpha}\right)\left((1-t)^{s}|f'(a)|+t^{s}\left|f'\left(\tfrac{a+b}{2}\right)\right|\right)dt\right.\\
&+\int_{\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}^{1}\left(t^{\alpha}-\tfrac{1}{4}\right)\left((1-t)^{s}|f'(a)|+t^{s}\left|f'\left(\tfrac{a+b}{2}\right)\right|\right)dt\\
&+\int_0^{1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}\left((1-t)^{\alpha}-\tfrac{1}{4}\right)\left((1-t)^{s}\left|f'\left(\tfrac{a+b}{2}\right)\right|+t^{s}|f'(b)|\right)dt\\
&\left.+\int_{1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}^{1}\left(\tfrac{1}{4}-(1-t)^{\alpha}\right)\left((1-t)^{s}\left|f'\left(\tfrac{a+b}{2}\right)\right|+t^{s}|f'(b)|\right)dt\right)\\
=\ & \tfrac{b-a}{4}\left(|f'(a)|\left(\int_0^{\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}\left(\tfrac{1}{4}-t^{\alpha}\right)(1-t)^{s}dt+\int_{\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}^{1}\left(t^{\alpha}-\tfrac{1}{4}\right)(1-t)^{s}dt\right)\right.\\
&+\left|f'\left(\tfrac{a+b}{2}\right)\right|\left(\int_0^{\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}\left(\tfrac{1}{4}-t^{\alpha}\right)t^{s}dt+\int_0^{1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}\left((1-t)^{\alpha}-\tfrac{1}{4}\right)(1-t)^{s}dt\right.
\end{aligned}
$$

$$+\int_{\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}^{1}\left(t^{\alpha}-\tfrac{1}{4}\right)t^{s}dt+\int_{1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}^{1}\left(\tfrac{1}{4}-(1-t)^{\alpha}\right)(1-t)^{s}dt\Bigg)$$

$$+\left|f'(b)\right|\left(\int_{0}^{1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}\left((1-t)^{\alpha}-\tfrac{1}{4}\right)t^{s}dt+\int_{1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}^{1}\left(\tfrac{1}{4}-(1-t)^{\alpha}\right)t^{s}dt\right)\Bigg)$$

$$=\frac{b-a}{4}\Bigg(\Theta_{s,\alpha}\left|f'(a)\right|$$

$$+\left(\frac{\alpha}{(s+1)(\alpha+s+1)}\left(\frac{1}{4}\right)^{\frac{s+1}{\alpha}}+\frac{3s+3-\alpha}{2(s+1)(\alpha+s+1)}\right)\left|f'\left(\frac{a+b}{2}\right)\right|+\Theta_{s,\alpha}\left|f'(b)\right|\Bigg),$$

where we have used

$$\int_{0}^{\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}\left(\tfrac{1}{4}-t^{\alpha}\right)(1-t)^{s}dt=\int_{1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}^{1}\left(\tfrac{1}{4}-(1-t)^{\alpha}\right)t^{s}dt \tag{5}$$

$$=\frac{1}{4(s+1)}\left(1-\left(1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}\right)^{s+1}\right)-B_{\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}(\alpha+1,s+1),$$

$$\int_{\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}^{1}\left(t^{\alpha}-\tfrac{1}{4}\right)(1-t)^{s}dt=\int_{0}^{1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}\left((1-t)^{\alpha}-\tfrac{1}{4}\right)t^{s}dt \tag{6}$$

$$=B_{1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}(s+1,\alpha+1)-\frac{1}{4(s+1)}\left(1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}\right)^{s+1},$$

$$\int_{0}^{\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}\left(\tfrac{1}{4}-t^{\alpha}\right)t^{s}dt=\int_{1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}^{1}\left(\tfrac{1}{4}-(1-t)^{\alpha}\right)(1-t)^{s}dt \tag{7}$$

$$=\frac{\alpha}{4(s+1)(\alpha+s+1)}\left(\frac{1}{4}\right)^{\frac{s+1}{\alpha}},$$

$$\int_{\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}^{1}\left(t^{\alpha}-\tfrac{1}{4}\right)t^{s}dt=\int_{0}^{1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}\left((1-t)^{\alpha}-\tfrac{1}{4}\right)(1-t)^{s}dt \tag{8}$$

$$=\frac{\alpha}{4(s+1)(\alpha+s+1)}\left(\frac{1}{4}\right)^{\frac{s+1}{\alpha}}+\frac{3s+3-\alpha}{4(s+1)(\alpha+s+1)},$$

and $\Theta_{s,\alpha}$ is defined as in (4). The proof is completed.

COROLLARY 2.6. In Theorem 2.5, if we take $s=1$, then we get

$$\left|\frac{1}{8}\left(f(a)+6f\left(\frac{a+b}{2}\right)+f(b)\right)-\frac{\Gamma(\alpha+1)}{2^{1-\alpha}(b-a)^{\alpha}}\left(I^{\alpha}_{\left(\frac{a+b}{2}\right)^{-}}f(a)+I^{\alpha}_{\left(\frac{a+b}{2}\right)^{+}}f(b)\right)\right|$$

$$\leq \tfrac{b-a}{4}\left(\Theta_{1,\alpha}\left|f'(a)\right|+\left(\tfrac{\alpha}{2(\alpha+2)}\left(\tfrac{1}{4}\right)^{\frac{2}{\alpha}}+\tfrac{6-\alpha}{4(\alpha+2)}\right)\left|f'\left(\tfrac{a+b}{2}\right)\right|+\Theta_{1,\alpha}\left|f'(b)\right|\right),$$

where

$$\Theta_{1,\alpha}=\tfrac{(\alpha+1)(\alpha+2)+8}{8(\alpha+1)(\alpha+2)}-\tfrac{1}{4}\left(1-\left(\tfrac{1}{4}\right)^{\frac{1}{\alpha}}\right)^2+2\left(\tfrac{(\alpha+1)-(\alpha+2)4^{\frac{1}{\alpha}}}{(\alpha+1)(\alpha+2)}\right)\left(\tfrac{1}{4}\right)^{\frac{\alpha+2}{\alpha}}. \tag{9}$$

COROLLARY 2.7. In Theorem 2.5, if we take $\alpha=1$, then we get

$$\begin{aligned}&\left|\tfrac{1}{8}\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right)-\tfrac{1}{b-a}\int_a^b f(t)\,dt\right|\\ \leq\ &\tfrac{b-a}{4(s+1)(s+2)}\left(\left(\tfrac{s-2}{4}+2\left(\tfrac{3}{4}\right)^{s+2}\right)\left|f'(a)\right|+\left(\left(\tfrac{1}{4}\right)^{s+1}+\tfrac{3s+2}{2}\right)\left|f'\left(\tfrac{a+b}{2}\right)\right|\right.\\ &\left.+\left(\tfrac{s-2}{4}+2\left(\tfrac{3}{4}\right)^{s+2}\right)\left|f'(b)\right|\right).\end{aligned}$$

Remark 2.8. Corollary 2.7 will be reduced to Corollary 3.1.3 from [17], if we take $s=1$.

THEOREM 2.9. Let $f:[a,b]\to\mathbb{R}$ be a differentiable function on (a,b) such that $f'\in L^1[a,b]$ with $0\leq a<b$. If $|f'|^q$ is s-convex in the second sense for some fixed $s\in(0,1]$ where $q>1$ with $\frac{1}{p}+\frac{1}{q}=1$, then we have

$$\begin{aligned}&\left|\tfrac{1}{8}\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right)-\tfrac{\Gamma(\alpha+1)}{2^{1-\alpha}(b-a)^\alpha}\left(I^\alpha_{\left(\frac{a+b}{2}\right)^-}f(a)+I^\alpha_{\left(\frac{a+b}{2}\right)^+}f(b)\right)\right|\\ \leq\ &\tfrac{b-a}{4}\left(\tfrac{1}{4^{p+\frac{1}{\alpha}}\alpha}B\left(\tfrac{1}{\alpha},p+1\right)+\tfrac{3^{p+1}}{4^{p+1}\alpha(p+1)}\cdot{}_2F_1\left(1-\tfrac{1}{\alpha},1,p+2;\tfrac{3}{4}\right)\right)^{\frac{1}{p}}\\ &\times\left(\left(\tfrac{\left|f'(a)\right|^q+\left|f'\left(\frac{a+b}{2}\right)\right|^q}{s+1}\right)^{\frac{1}{q}}+\left(\tfrac{\left|f'\left(\frac{a+b}{2}\right)\right|^q+\left|f'(b)\right|^q}{s+1}\right)^{\frac{1}{q}}\right).\end{aligned}$$

PROOF. From Lemma 2.1, properties of modulus, Hölder's inequality, and s-convexity in the second sense of $|f'|^q$, we have

$$\begin{aligned}&\left|\tfrac{1}{8}\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right)-\tfrac{\Gamma(\alpha+1)}{2^{1-\alpha}(b-a)^\alpha}\left(I^\alpha_{\left(\frac{a+b}{2}\right)^-}f(a)+I^\alpha_{\left(\frac{a+b}{2}\right)^+}f(b)\right)\right|\\ \leq\ &\tfrac{b-a}{4}\left(\left(\int_0^1\left|t^\alpha-\tfrac{1}{4}\right|^p dt\right)^{\frac{1}{p}}\left(\int_0^1\left|f'\left((1-t)a+t\tfrac{a+b}{2}\right)\right|^q dt\right)^{\frac{1}{q}}\right.\\ &\left.+\left(\int_0^1\left|(1-t)^\alpha-\tfrac{1}{4}\right|^p dt\right)^{\frac{1}{p}}\left(\int_0^1\left|f'\left((1-t)\tfrac{a+b}{2}+tb\right)\right|^q dt\right)^{\frac{1}{q}}\right)\\ \leq\ &\tfrac{b-a}{4}\left(\left(\int_0^{\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}\left(\tfrac{1}{4}-t^\alpha\right)^p dt+\int_{\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}^1\left(t^\alpha-\tfrac{1}{4}\right)^p dt\right)^{\frac{1}{p}}\right.\end{aligned}$$

$$\times\left(\int_0^1\left|f'\left((1-t)a+t\tfrac{a+b}{2}\right)\right|^q dt\right)^{\frac{1}{q}}$$

$$+\left(\int_0^{1-\left(\frac14\right)^{\frac1\alpha}}\left((1-t)^\alpha-\tfrac14\right)^p dt+\int_{1-\left(\frac14\right)^{\frac1\alpha}}^1\left(\tfrac14-(1-t)^\alpha\right)^p dt\right)^{\frac1p}$$

$$\left.\times\left(\int_0^1\left|f'\left((1-t)\tfrac{a+b}{2}+tb\right)\right|^q dt\right)^{\frac1q}\right)$$

$$\leq \frac{b-a}{4}\left(\int_0^{\left(\frac14\right)^{\frac1\alpha}}\left(\tfrac14-t^\alpha\right)^p dt+\int_{\left(\frac14\right)^{\frac1\alpha}}^1\left(t^\alpha-\tfrac14\right)^p dt\right)^{\frac1p}$$

$$\times\left(\left(\int_0^1\left((1-t)^s\left|f'(a)\right|^q+t^s\left|f'\left(\tfrac{a+b}{2}\right)\right|^q\right)dt\right)^{\frac1q}\right.$$

$$\left.+\left(\int_0^1\left((1-t)^s\left|f'\left(\tfrac{a+b}{2}\right)\right|^q+t^s\left|f'(b)\right|^q\right)dt\right)^{\frac1q}\right)$$

$$= \frac{b-a}{4}\left(\frac{1}{4^{p+\frac1\alpha}\alpha}B\left(\tfrac1\alpha,p+1\right)+\frac{3^{p+1}}{4^{p+1}\alpha(p+1)}\cdot{}_2F_1\left(1-\tfrac1\alpha,1,p+2;\tfrac34\right)\right)^{\frac1p}$$

$$\times\left(\left(\frac{\left|f'(a)\right|^q+\left|f'\left(\frac{a+b}{2}\right)\right|^q}{s+1}\right)^{\frac1q}+\left(\frac{\left|f'\left(\frac{a+b}{2}\right)\right|^q+\left|f'(b)\right|^q}{s+1}\right)^{\frac1q}\right),$$

where we have used the fact that

$$\int_0^{\left(\frac14\right)^{\frac1\alpha}}\left(\tfrac14-t^\alpha\right)^p dt=\frac1\alpha\int_0^{\frac14}\left(\tfrac14-w\right)^p w^{\frac1\alpha-1}dw=\frac{1}{4^p\alpha}\int_0^{\frac14}(1-4w)^p w^{\frac1\alpha-1}dw$$

$$=\frac{1}{4^{p+\frac1\alpha}\alpha}\int_0^1 u^{\frac1\alpha-1}(1-u)^p du=\frac{1}{4^{p+\frac1\alpha}\alpha}B\left(\tfrac1\alpha,p+1\right),$$

and

$$\int_{\left(\frac14\right)^{\frac1\alpha}}^1\left(t^\alpha-\tfrac14\right)^p dt=\frac1\alpha\int_{\frac14}^1\left(w-\tfrac14\right)^p w^{\frac1\alpha-1}dw=\left(\tfrac34\right)^p\frac1\alpha\int_0^{\frac34}\left(1-\tfrac43u\right)^p(1-u)^{\frac1\alpha-1}du$$

$$=\left(\tfrac34\right)^{p+1}\frac1\alpha\int_0^1(1-z)^p\left(1-\tfrac34z\right)^{\frac1\alpha-1}dz$$

$$=\frac{3^{p+1}}{4^{p+1}\alpha(p+1)}\cdot{}_2F_1\left(1-\tfrac1\alpha,1,p+2;\tfrac34\right).$$

The proof is completed.

COROLLARY 2.10. In Theorem 2.9, if we take $s=1$, then we get

$$\left|\frac18\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right)-\frac{\Gamma(\alpha+1)}{2^{1-\alpha}(b-a)^\alpha}\left(I^\alpha_{\left(\frac{a+b}{2}\right)^-}f(a)+I^\alpha_{\left(\frac{a+b}{2}\right)^+}f(b)\right)\right|$$

$$\leq \frac{b-a}{4}\left(\frac{1}{4^{p+\frac{1}{\alpha}}\alpha}B\left(\frac{1}{\alpha},p+1\right)+\frac{3^{p+1}}{4^{p+1}\alpha(p+1)}\cdot{}_2F_1\left(1-\frac{1}{\alpha},1,p+2;\frac{3}{4}\right)\right)^{\frac{1}{p}}$$
$$\times\left(\left(\frac{\left|f'(a)\right|^q+\left|f'\left(\frac{a+b}{2}\right)\right|^q}{2}\right)^{\frac{1}{q}}+\left(\frac{\left|f'\left(\frac{a+b}{2}\right)\right|^q+\left|f'(b)\right|^q}{2}\right)^{\frac{1}{q}}\right).$$

COROLLARY 2.11. In Theorem 2.9, if we take $\alpha = 1$, then we get

$$\left|\frac{1}{8}\left(f(a)+6f\left(\frac{a+b}{2}\right)+f(b)\right)-\frac{1}{b-a}\int_a^b f(t)\,dt\right|$$
$$\leq \frac{b-a}{16}\left(\frac{1+3^{p+1}}{4(p+1)}\right)^{\frac{1}{p}}\left(\left(\frac{\left|f'(a)\right|^q+\left|f'\left(\frac{a+b}{2}\right)\right|^q}{s+1}\right)^{\frac{1}{q}}+\left(\frac{\left|f'\left(\frac{a+b}{2}\right)\right|^q+\left|f'(b)\right|^q}{s+1}\right)^{\frac{1}{q}}\right).$$

Remark 2.12. Corollary 2.11 will be reduced to Corollary 3.2.1 from [17], if we take $s = 1$.

THEOREM 2.13. Let $f : [a,b] \to \mathbb{R}$ be a differentiable function on (a,b) such that $f' \in L^1[a,b]$ with $0 \leq a < b$. If $|f'|^q$ is s-convex in the second sense for some fixed $s \in (0,1]$ where $q \geq 1$, then we have

$$\left|\frac{1}{8}\left(f(a)+6f\left(\frac{a+b}{2}\right)+f(b)\right)-\frac{\Gamma(\alpha+1)}{2^{1-\alpha}(b-a)^{\alpha}}\left(I^{\alpha}_{\left(\frac{a+b}{2}\right)^-}f(a)+I^{\alpha}_{\left(\frac{a+b}{2}\right)^+}f(b)\right)\right|$$
$$\leq \frac{b-a}{4}\left(\frac{\left(2-4^{\frac{1}{\alpha}}\right)\alpha+3\times 4^{\frac{1}{\alpha}}}{4^{\frac{\alpha+1}{\alpha}}(\alpha+1)}\right)^{1-\frac{1}{q}}$$
$$\times\left(\left(\Theta_{s,\alpha}\left|f'(a)\right|^q+\frac{\left(2-4^{\frac{s+1}{\alpha}}\right)\alpha+(3s+3)4^{\frac{s+1}{\alpha}}}{4^{\frac{s+\alpha+1}{\alpha}}(s+1)(\alpha+s+1)}\left|f'\left(\frac{a+b}{2}\right)\right|^q\right)^{\frac{1}{q}}\right.$$
$$\left.+\left(\frac{\left(2-4^{\frac{s+1}{\alpha}}\right)\alpha+(3s+3)4^{\frac{s+1}{\alpha}}}{4^{\frac{s+\alpha+1}{\alpha}}(s+1)(\alpha+s+1)}\left|f'\left(\frac{a+b}{2}\right)\right|^q+\Theta_{s,\alpha}\left|f'(b)\right|^q\right)^{\frac{1}{q}}\right),$$

where $\Theta_{s,\alpha}$ is defined as in (4).

PROOF. From Lemma 2.1, properties of modulus, power mean inequality, and s-convexity in the second sense of $|f'|^q$, we have

$$\left|\frac{1}{8}\left(f(a)+6f\left(\frac{a+b}{2}\right)+f(b)\right)-\frac{\Gamma(\alpha+1)}{2^{1-\alpha}(b-a)^{\alpha}}\left(I^{\alpha}_{\left(\frac{a+b}{2}\right)^-}f(a)+I^{\alpha}_{\left(\frac{a+b}{2}\right)^+}f(b)\right)\right|$$
$$\leq \frac{b-a}{4}\left(\left(\int_0^1\left|t^{\alpha}-\frac{1}{4}\right|dt\right)^{1-\frac{1}{q}}\left(\int_0^1\left|t^{\alpha}-\frac{1}{4}\right|\left|f'\left((1-t)a+t\frac{a+b}{2}\right)\right|^q dt\right)^{\frac{1}{q}}\right.$$
$$+\left(\int_0^1\left|(1-t)^{\alpha}-\frac{1}{4}\right|dt\right)^{1-\frac{1}{q}}$$

$$\times\left(\int_0^1\left|(1-t)^\alpha-\tfrac14\right|\left|f'\left((1-t)\tfrac{a+b}{2}+tb\right)\right|^q dt\right)^{\frac1q}\Bigg)$$

$$\leq \frac{b-a}{4}\left(\left(\int_0^{\left(\frac14\right)^{\frac1\alpha}}\left(\tfrac14-t^\alpha\right)dt+\int_{\left(\frac14\right)^{\frac1\alpha}}^1\left(t^\alpha-\tfrac14\right)dt\right)^{1-\frac1q}\right.$$

$$\times\left(\int_0^1\left|t^\alpha-\tfrac14\right|\left|f'\left((1-t)a+t\tfrac{a+b}{2}\right)\right|^q dt\right)^{\frac1q}$$

$$+\left(\int_0^{1-\left(\frac14\right)^{\frac1\alpha}}\left((1-t)^\alpha-\tfrac14\right)dt+\int_{1-\left(\frac14\right)^{\frac1\alpha}}^1\left(\tfrac14-(1-t)^\alpha\right)dt\right)^{1-\frac1q}$$

$$\times\left(\int_0^1\left|(1-t)^\alpha-\tfrac14\right|\left|f'\left((1-t)\tfrac{a+b}{2}+tb\right)\right|^q dt\right)^{\frac1q}\Bigg)$$

$$\leq \frac{b-a}{4}\left(\int_0^{\left(\frac14\right)^{\frac1\alpha}}\left(\tfrac14-t^\alpha\right)dt+\int_{\left(\frac14\right)^{\frac1\alpha}}^1\left(t^\alpha-\tfrac14\right)dt\right)^{1-\frac1q}$$

$$\times\left(\left(\int_0^1\left|t^\alpha-\tfrac14\right|\left((1-t)^s\left|f'(a)\right|^q+t^s\left|f'\left(\tfrac{a+b}{2}\right)\right|^q\right)dt\right)^{\frac1q}\right.$$

$$+\left.\left(\int_0^1\left|(1-t)^\alpha-\tfrac14\right|\left((1-t)^s\left|f'\left(\tfrac{a+b}{2}\right)\right|^q+t^s\left|f'(b)\right|^q\right)dt\right)^{\frac1q}\right)$$

$$= \frac{b-a}{4}\left(\frac{\left(2-4^{\frac1\alpha}\right)\alpha+3\times4^{\frac1\alpha}}{4^{\frac{\alpha+1}{\alpha}}(\alpha+1)}\right)^{1-\frac1q}$$

$$\times\left(\left(\left|f'(a)\right|^q\left(\int_0^{\left(\frac14\right)^{\frac1\alpha}}\left(\tfrac14-t^\alpha\right)(1-t)^s dt+\int_{\left(\frac14\right)^{\frac1\alpha}}^1\left(t^\alpha-\tfrac14\right)(1-t)^s dt\right)\right.\right.$$

$$\left.\left.+\left|f'\left(\tfrac{a+b}{2}\right)\right|^q\left(\int_0^{\left(\frac14\right)^{\frac1\alpha}}\left(\tfrac14-t^\alpha\right)t^s dt+\int_{\left(\frac14\right)^{\frac1\alpha}}^1\left(t^\alpha-\tfrac14\right)t^s dt\right)\right)\right)^{\frac1q}$$

$$+\left(\left(\int_0^{1-\left(\frac14\right)^{\frac1\alpha}}\left((1-t)^\alpha-\tfrac14\right)(1-t)^s dt\right.\right.$$

$$\left.+\int_{1-\left(\frac14\right)^{\frac1\alpha}}^1\left(\tfrac14-(1-t)^\alpha\right)(1-t)^s dt\right)\left|f'\left(\tfrac{a+b}{2}\right)\right|^q$$

$$+\left(\int\limits_{0}^{1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}\left((1-t)^{\alpha}-\tfrac{1}{4}\right)t^{s}dt+\int\limits_{1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}^{1}\left(\tfrac{1}{4}-(1-t)^{\alpha}\right)t^{s}dt\right)\left|f'(b)\right|^{q}\Bigg)^{\frac{1}{q}}\Bigg)$$

$$= \frac{b-a}{4}\left(\frac{\left(2-4^{\frac{1}{\alpha}}\right)\alpha+3\times4^{\frac{1}{\alpha}}}{4^{\frac{\alpha+1}{\alpha}}(\alpha+1)}\right)^{1-\frac{1}{q}}$$

$$\times\left(\left(\Theta_{s,\alpha}\left|f'(a)\right|^{q}+\frac{\left(2-4^{\frac{s+1}{\alpha}}\right)\alpha+(3s+3)4^{\frac{s+1}{\alpha}}}{4^{\frac{s+\alpha+1}{\alpha}}(s+1)(\alpha+s+1)}\left|f'\left(\tfrac{a+b}{2}\right)\right|^{q}\right)^{\frac{1}{q}}\right.$$

$$\left.+\left(\frac{\left(2-4^{\frac{s+1}{\alpha}}\right)\alpha+(3s+3)4^{\frac{s+1}{\alpha}}}{4^{\frac{s+\alpha+1}{\alpha}}(s+1)(\alpha+s+1)}\left|f'\left(\tfrac{a+b}{2}\right)\right|^{q}+\Theta_{s,\alpha}\left|f'(b)\right|^{q}\right)^{\frac{1}{q}}\right),$$

where we have used (4)-(8). The proof is achieved.

COROLLARY 2.14. In Theorem 2.13, if we take $s=1$, then we get

$$\left|\tfrac{1}{8}\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right)-\tfrac{\Gamma(\alpha+1)}{2^{1-\alpha}(b-a)^{\alpha}}\left(I^{\alpha}_{\left(\frac{a+b}{2}\right)^{-}}f(a)+I^{\alpha}_{\left(\frac{a+b}{2}\right)^{+}}f(b)\right)\right|$$

$$\leq \frac{b-a}{4}\left(\frac{\left(2-4^{\frac{1}{\alpha}}\right)\alpha+3\times4^{\frac{1}{\alpha}}}{4^{\frac{\alpha+1}{\alpha}}(\alpha+1)}\right)^{1-\frac{1}{q}}$$

$$\times\left(\left(\Theta_{1,\alpha}\left|f'(a)\right|^{q}+\frac{\left(2-4^{\frac{2}{\alpha}}\right)\alpha+6\times4^{\frac{2}{\alpha}}}{4^{\frac{2+3\alpha}{\alpha}}(\alpha+2)}\left|f'\left(\tfrac{a+b}{2}\right)\right|^{q}\right)^{\frac{1}{q}}\right.$$

$$\left.+\left(\frac{\left(2-4^{\frac{2}{\alpha}}\right)\alpha+6\times4^{\frac{2}{\alpha}}}{4^{\frac{2+3\alpha}{\alpha}}(\alpha+2)}\left|f'\left(\tfrac{a+b}{2}\right)\right|^{q}+\Theta_{1,\alpha}\left|f'(b)\right|^{q}\right)^{\frac{1}{q}}\right),$$

where $\Theta_{1,\alpha}$ is defined in (9).

COROLLARY 2.15. In Theorem 2.13, if we take $\alpha=1$, then we get

$$\left|\tfrac{1}{8}\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right)-\tfrac{1}{b-a}\int\limits_{a}^{b}f(t)\,dt\right|$$

$$\leq \frac{b-a}{4[(s+1)(s+2)]^{\frac{1}{q}}}\left(\tfrac{5}{16}\right)^{1-\frac{1}{q}}$$

$$\times\left(\left(\left(\tfrac{s-2}{4}+2\left(\tfrac{3}{4}\right)^{s+2}\right)\left|f'(a)\right|^{q}+\tfrac{2+(3s+2)4^{s+1}}{4^{s+2}}\left|f'\left(\tfrac{a+b}{2}\right)\right|^{q}\right)^{\frac{1}{q}}\right.$$

$$\left.+\left(\tfrac{2+(3s+2)4^{s+1}}{4^{s+2}}\left|f'\left(\tfrac{a+b}{2}\right)\right|^{q}+\left(\tfrac{s-2}{4}+2\left(\tfrac{3}{4}\right)^{s+2}\right)\left|f'(b)\right|^{q}\right)^{\frac{1}{q}}\right).$$

Remark 2.16. Corollary 2.15 will be reduced to Corollary 3.1.2 from [17], if we take $s=1$.

3. FURTHER RESULTS

THEOREM 3.1. Let $f:[a,b]\to\mathbb{R}$ be a differentiable function on (a,b) such that $f'\in L^1[a,b]$ with $0\le a<b$. If there exist constants $-\infty<m<M<+\infty$ such that $m\le f'(x)\le M$ for all $x\in[a,b]$, then we have

$$\begin{aligned}&\left|\tfrac{1}{8}\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right)-\tfrac{\Gamma(\alpha+1)}{2^{1-\alpha}(b-a)^\alpha}\left(I^\alpha_{\left(\frac{a+b}{2}\right)^-}f(a)+I^\alpha_{\left(\frac{a+b}{2}\right)^+}f(b)\right)\right|\\ \le\;& \tfrac{(b-a)(M-m)}{4(\alpha+1)}\left(\tfrac{\alpha}{2}\left(\tfrac{1}{4}\right)^{\frac{1}{\alpha}}+\tfrac{3-\alpha}{4}\right).\end{aligned}$$

PROOF. From Lemma 2.1, we have

$$\begin{aligned}&\tfrac{1}{8}\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right)-\tfrac{\Gamma(\alpha+1)}{2^{1-\alpha}(b-a)^\alpha}\left(I^\alpha_{\left(\frac{a+b}{2}\right)^-}f(a)+I^\alpha_{\left(\frac{a+b}{2}\right)^+}f(b)\right)\\
=\;&\tfrac{b-a}{4}\left(\int_0^1\left(t^\alpha-\tfrac{1}{4}\right)f'\left((1-t)a+t\tfrac{a+b}{2}\right)dt\right.\\
&\left.-\int_0^1\left((1-t)^\alpha-\tfrac{1}{4}\right)f'\left((1-t)\tfrac{a+b}{2}+tb\right)dt\right)\\
=\;&\tfrac{b-a}{4}\left(\int_0^1\left(t^\alpha-\tfrac{1}{4}\right)\left(f'\left((1-t)a+t\tfrac{a+b}{2}\right)-\tfrac{m+M}{2}+\tfrac{m+M}{2}\right)dt\right.\\
&\left.-\int_0^1\left((1-t)^\alpha-\tfrac{1}{4}\right)\left(f'\left((1-t)\tfrac{a+b}{2}+tb\right)-\tfrac{m+M}{2}+\tfrac{m+M}{2}\right)dt\right)\\
=\;&\tfrac{b-a}{4}\left(\int_0^1\left(t^\alpha-\tfrac{1}{4}\right)\left(f'\left((1-t)a+t\tfrac{a+b}{2}\right)-\tfrac{m+M}{2}\right)dt\right.\\
&-\int_0^1\left((1-t)^\alpha-\tfrac{1}{4}\right)\left(f'\left((1-t)\tfrac{a+b}{2}+tb\right)-\tfrac{m+M}{2}\right)dt\\
&\left.+\tfrac{m+M}{2}\left(\int_0^1\left(t^\alpha-\tfrac{1}{4}\right)dt-\int_0^1\left((1-t)^\alpha-\tfrac{1}{4}\right)dt\right)\right)\\
=\;&\tfrac{b-a}{4}\left(\int_0^1\left(t^\alpha-\tfrac{1}{4}\right)\left(f'\left((1-t)a+t\tfrac{a+b}{2}\right)-\tfrac{m+M}{2}\right)dt\right.\\
&\left.-\int_0^1\left((1-t)^\alpha-\tfrac{1}{4}\right)\left(f'\left((1-t)\tfrac{a+b}{2}+tb\right)-\tfrac{m+M}{2}\right)dt\right),\end{aligned}\tag{10}$$

where we have taken into consideration that

$$\int_0^1\left(t^\alpha-\tfrac{1}{4}\right)dt-\int_0^1\left((1-t)^\alpha-\tfrac{1}{4}\right)dt=\left(\tfrac{1}{\alpha+1}-\tfrac{1}{4}\right)-\left(\tfrac{1}{\alpha+1}-\tfrac{1}{4}\right)=0.$$

Applying the absolute value to both sides of (10), we get

$$\begin{aligned}&\left|\tfrac{1}{8}\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right)-\tfrac{\Gamma(\alpha+1)}{2^{1-\alpha}(b-a)^\alpha}\left(I^\alpha_{\left(\frac{a+b}{2}\right)^-}f(a)+I^\alpha_{\left(\frac{a+b}{2}\right)^+}f(b)\right)\right|\\
=\;&\tfrac{b-a}{4}\left(\int_0^1\left|t^\alpha-\tfrac{1}{4}\right|\left|f'\left((1-t)a+t\tfrac{a+b}{2}\right)-\tfrac{m+M}{2}\right|dt\right.\end{aligned}$$

$$+\int_0^1 \left|(1-t)^{\alpha}-\tfrac{1}{4}\right|\left|f'\left((1-t)\tfrac{a+b}{2}+tb\right)-\tfrac{m+M}{2}\right|dt\Bigg). \tag{11}$$

Since $m \le f'(x) \le M$ for all $x \in [a,b]$, we have

$$\left|f'\left((1-t)a+t\tfrac{a+b}{2}\right)-\tfrac{m+M}{2}\right| \le \tfrac{M-m}{2}, \tag{12}$$

and

$$\left|f'\left((1-t)\tfrac{a+b}{2}+tb\right)-\tfrac{m+M}{2}\right| \le \tfrac{M-m}{2}. \tag{13}$$

Using (12) and (13) in (11), we get

$$\begin{aligned}
&\left|\tfrac{1}{8}\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right)-\tfrac{\Gamma(\alpha+1)}{2^{1-\alpha}(b-a)^{\alpha}}\left(I^{\alpha}_{\left(\frac{a+b}{2}\right)^-}f(a)+I^{\alpha}_{\left(\frac{a+b}{2}\right)^+}f(b)\right)\right|\\
= \;& \tfrac{(b-a)(M-m)}{8}\left(\int_0^1\left|t^{\alpha}-\tfrac{1}{4}\right|dt+\int_0^1\left|(1-t)^{\alpha}-\tfrac{1}{4}\right|dt\right)\\
= \;& \tfrac{(b-a)(M-m)}{4}\left(\int_0^1\left|t^{\alpha}-\tfrac{1}{4}\right|dt\right)\\
= \;& \tfrac{(b-a)(M-m)}{4(\alpha+1)}\left(\tfrac{\alpha}{2}\left(\tfrac{1}{4}\right)^{\frac{1}{\alpha}}+\tfrac{3-\alpha}{4}\right).
\end{aligned}$$

The proof is completed.

COROLLARY 3.2. In Theorem 3.1, if we take $\alpha = 1$, then we get

$$\left|\tfrac{1}{8}\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right)-\tfrac{1}{b-a}\int_a^b f(t)\,dt\right| \le \tfrac{5(b-a)(M-m)}{64}.$$

Remark 3.3. The result of Corollary 2.11 is finer than the result of Corollary 3.1 of [14].

THEOREM 3.4. Let $f:[a,b]\to\mathbb{R}$ be a differentiable function on (a,b) such that $f' \in L^1[a,b]$ with $0 \le a < b$. If f' is r-L-Hölderian function on $[a,b]$ (i.e. there exist $L > 0$ and $0 < r \le 1$ such that $|f'(x)-f'(y)| \le L|x-y|^r$), then we have

$$\begin{aligned}
&\left|\tfrac{1}{8}\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right)-\tfrac{\Gamma(\alpha+1)}{2^{1-\alpha}(b-a)^{\alpha}}\left(I^{\alpha}_{\left(\frac{a+b}{2}\right)^-}f(a)+I^{\alpha}_{\left(\frac{a+b}{2}\right)^+}f(b)\right)\right|\\
\le \;& \tfrac{L(b-a)^{r+1}}{2^{r+2}}\Bigg(\tfrac{-\alpha^2+(6-r)\alpha+3r+7}{4(\alpha+1)(\alpha+r+1)}+\tfrac{\alpha}{2(r+1)(\alpha+r+1)}\left(\tfrac{1}{4}\right)^{\frac{r+1}{\alpha}}-\tfrac{1}{2(r+1)}\left(1-\left(\tfrac{1}{4}\right)^{\frac{1}{\alpha}}\right)^{r+1}\\
&+B_{1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}(r+1,\alpha+1)-B_{\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}(\alpha+1,r+1)\Bigg).
\end{aligned}$$

PROOF. From Lemma 2.1, we have

$$\tfrac{1}{8}\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right)-\tfrac{\Gamma(\alpha+1)}{2^{1-\alpha}(b-a)^{\alpha}}\left(I^{\alpha}_{\left(\frac{a+b}{2}\right)^-}f(a)+I^{\alpha}_{\left(\frac{a+b}{2}\right)^+}f(b)\right)$$

$$
\begin{aligned}
= \quad & \tfrac{b-a}{4}\left(\int_0^1 \left(t^{\alpha}-\tfrac{1}{4}\right) f'\left((1-t)a+t\tfrac{a+b}{2}\right)dt\right. \\
& \left. -\int_0^1 \left((1-t)^{\alpha}-\tfrac{1}{4}\right) f'\left((1-t)\tfrac{a+b}{2}+tb\right)dt\right) \\
= \quad & \tfrac{b-a}{4}\left(\int_0^1 \left(t^{\alpha}-\tfrac{1}{4}\right)\left(f'\left((1-t)a+t\tfrac{a+b}{2}\right)-f'(a)+f'(a)\right)dt\right. \\
& \left. -\int_0^1 \left((1-t)^{\alpha}-\tfrac{1}{4}\right)\left(f'\left((1-t)\tfrac{a+b}{2}+tb\right)-f'\left(\tfrac{a+b}{2}\right)+f'\left(\tfrac{a+b}{2}\right)\right)dt\right) \\
= \quad & \tfrac{b-a}{4}\left(\int_0^1 \left(t^{\alpha}-\tfrac{1}{4}\right)\left(f'\left((1-t)a+t\tfrac{a+b}{2}\right)-f'(a)\right)dt\right. \\
& -\int_0^1 \left((1-t)^{\alpha}-\tfrac{1}{4}\right)\left(f'\left((1-t)\tfrac{a+b}{2}+tb\right)-f'\left(\tfrac{a+b}{2}\right)\right)dt \\
& \left. +f'(a)\int_0^1 \left(t^{\alpha}-\tfrac{1}{4}\right)dt - f'\left(\tfrac{a+b}{2}\right)\int_0^1 \left((1-t)^{\alpha}-\tfrac{1}{4}\right)dt\right) \\
= \quad & \tfrac{b-a}{4}\left(\int_0^1 \left(t^{\alpha}-\tfrac{1}{4}\right)\left(f'\left((1-t)a+t\tfrac{a+b}{2}\right)-f'(a)\right)dt\right. \\
& -\int_0^1 \left((1-t)^{\alpha}-\tfrac{1}{4}\right)\left(f'\left((1-t)\tfrac{a+b}{2}+tb\right)-f'\left(\tfrac{a+b}{2}\right)\right)dt \\
& \left. +\tfrac{3-\alpha}{4(\alpha+1)}\left(f'(a)-f'\left(\tfrac{a+b}{2}\right)\right)\right). \qquad (14)
\end{aligned}
$$

Applying the absolute value in both sides of (14), and by using the fact that f' is r-L-Hölderian on $[a,b]$, we obtain

$$
\begin{aligned}
& \left|\tfrac{1}{8}\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right)-\tfrac{\Gamma(\alpha+1)}{2^{1-\alpha}(b-a)^{\alpha}}\left(I^{\alpha}_{\left(\frac{a+b}{2}\right)^-}f(a)+I^{\alpha}_{\left(\frac{a+b}{2}\right)^+}f(b)\right)\right| \\
\leq \quad & \tfrac{b-a}{4}\left(\int_0^1 \left|t^{\alpha}-\tfrac{1}{4}\right|\left|f'\left((1-t)a+t\tfrac{a+b}{2}\right)-f'(a)\right|dt\right. \\
& +\int_0^1 \left|(1-t)^{\alpha}-\tfrac{1}{4}\right|\left|f'\left((1-t)\tfrac{a+b}{2}+tb\right)-f'\left(\tfrac{a+b}{2}\right)\right|dt \\
& \left. +\tfrac{3-\alpha}{4(\alpha+1)}\left(\left|f'(a)-f'\left(\tfrac{a+b}{2}\right)\right|\right)\right) \\
\leq \quad & \tfrac{b-a}{4}L\left(\int_0^1 \left|t^{\alpha}-\tfrac{1}{4}\right|\left|(1-t)a+t\tfrac{a+b}{2}-a\right|^r dt\right. \\
& \left. +\int_0^1 \left|(1-t)^{\alpha}-\tfrac{1}{4}\right|\left|(1-t)\tfrac{a+b}{2}+tb-\tfrac{a+b}{2}\right|^r dt+\tfrac{3-\alpha}{4(\alpha+1)}\left(\left|a-\tfrac{a+b}{2}\right|^r\right)\right) \\
= \quad & \tfrac{L}{2}\left(\tfrac{b-a}{2}\right)^{r+1}\left(\int_0^1 \left|t^{\alpha}-\tfrac{1}{4}\right|t^r dt+\int_0^1 \left|(1-t)^{\alpha}-\tfrac{1}{4}\right|t^r dt+\tfrac{3-\alpha}{4(\alpha+1)}\right) \\
= \quad & \tfrac{L(b-a)^{r+1}}{2^{r+2}}\left(\tfrac{-\alpha^2+(6-r)\alpha+3r+7}{4(\alpha+1)(\alpha+r+1)}+\tfrac{\alpha}{2(r+1)(\alpha+r+1)}\left(\tfrac{1}{4}\right)^{\frac{r+1}{\alpha}}-\tfrac{1}{2(r+1)}\left(1-\left(\tfrac{1}{4}\right)^{\frac{1}{\alpha}}\right)^{r+1}\right. \\
& \left. B_{1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}(r+1,\alpha+1)-B_{\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}(\alpha+1,r+1)\right),
\end{aligned}
$$

where we have used

$$\int_0^1 \left|t^\alpha - \tfrac{1}{4}\right| t^r dt = \tfrac{\alpha}{2(r+1)(\alpha+r+1)} \left(\tfrac{1}{4}\right)^{\frac{r+1}{\alpha}} + \tfrac{3r+3-\alpha}{4(r+1)(\alpha+r+1)}$$

and

$$\int_0^1 \left|(1-t)^\alpha - \tfrac{1}{4}\right| t^r dt = B_{1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}(r+1,\alpha+1) - B_{\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}(\alpha+1,r+1)$$
$$+\tfrac{1}{4(r+1)} - \tfrac{1}{2(r+1)} \left(1-\left(\tfrac{1}{4}\right)^{\frac{1}{\alpha}}\right)^{r+1}.$$

The proof is completed.

Remark 3.5. By a simple calculation of definite integrals, we have

1/ $B_{1-\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}(2,\alpha+1) = \frac{1}{(\alpha+1)(\alpha+2)} - \frac{1}{4(\alpha+1)}\left(\frac{1}{4}\right)^{\frac{1}{\alpha}} + \frac{1}{4(\alpha+2)}\left(\frac{1}{4}\right)^{\frac{2}{\alpha}}$.

2/ $B_{\left(\frac{1}{4}\right)^{\frac{1}{\alpha}}}(\alpha+1,2) = \frac{1}{4(\alpha+1)}\left(\frac{1}{4}\right)^{\frac{1}{\alpha}} - \frac{1}{4(\alpha+2)}\left(\frac{1}{4}\right)^{\frac{2}{\alpha}}$.

3/ $B_{\frac{3}{4}}(r+1,2) = \frac{r+5}{4(r+1)(r+2)}\left(\frac{3}{4}\right)^{r+1}$.

4/ $B_{\frac{1}{4}}(2,r+1) = \frac{1}{(r+1)(r+2)} - \frac{r+5}{4(r+1)(r+2)}\left(\frac{3}{4}\right)^{r+1}$.

5/ $B_{\frac{3}{4}}(2,2) = \frac{9}{64}$.

6/ $B_{\frac{1}{4}}(2,2) = \frac{5}{192}$.

COROLLARY 3.6. Under the assumptions of Theorem 3.4, if f' is a Lipschitzian function, then we have

$$\left|\tfrac{1}{8}\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right) - \tfrac{\Gamma(\alpha+1)}{2^{1-\alpha}(b-a)^\alpha}\left(I^\alpha_{\left(\frac{a+b}{2}\right)^-} f(a) + I^\alpha_{\left(\frac{a+b}{2}\right)^+} f(b)\right)\right|$$
$$\leq \tfrac{L(b-a)^2}{8}\left(\tfrac{6+\alpha-\alpha^2}{2(\alpha+1)(\alpha+2)} + \tfrac{\alpha}{2(\alpha+1)}\left(\tfrac{1}{4}\right)^{\frac{1}{\alpha}}\right).$$

COROLLARY 3.7. In Theorem 3.4, if we take $\alpha = 1$, then we get

$$\left|\tfrac{1}{8}\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right) - \tfrac{1}{b-a}\int_a^b f(t)\,dt\right|$$
$$\leq \tfrac{L(b-a)^{r+1}}{2^{r+2}}\left(\tfrac{r^2+7r+2}{4(r+1)(r+2)} + \tfrac{2\left(1+3^{r+2}\right)}{4^{r+2}(r+1)(r+2)}\right).$$

COROLLARY 3.8. In Corollary 3.7, if we take $r = 1$, then we get

$$\left|\tfrac{1}{8}\left(f(a)+6f\left(\tfrac{a+b}{2}\right)+f(b)\right) - \tfrac{1}{b-a}\int_a^b f(t)\,dt\right| \leq \tfrac{9L(b-a)^2}{128}.$$

4. APPLICATIONS

Simpson like quadrature formula

Let Υ be the partition of the points $a = x_0 < x_1 < ... < x_n = b$ of the interval $[a,b]$, and consider the quadrature formula

$$\int_a^b f(u)\,du = \lambda(f,\Upsilon) + R(f,\Upsilon),$$

where

$$\lambda(f,\Upsilon) = \sum_{i=0}^{n-1} \tfrac{x_{i+1}-x_i}{8}\left(f(x_i) + 6f\left(\tfrac{x_i+x_{i+1}}{2}\right) + f(x_{i+1})\right)$$

and $R(f,\Upsilon)$ denotes the associated approximation error.

PROPOSITION 4.1. Let $n \in \mathbb{N}$ and $f : [a,b] \to \mathbb{R}$ be a differentiable function on (a,b) with $0 \le a < b$ and $f' \in L^1[a,b]$. If $|f'|$ is s-convex function in the second sense for some fixed $s \in (0,1]$, we have

$$\begin{aligned} |R(f,\Upsilon)| \le \quad & \sum_{i=0}^{n-1} \tfrac{(x_{i+1}-x_i)^2}{4(s+1)(s+2)} \left(\left(\tfrac{s-2}{4} + 2\left(\tfrac{3}{4}\right)^{s+2} \right) |f'(x_i)| \right. \\ & \left. + \left(\left(\tfrac{1}{4}\right)^{s+1} + \tfrac{3s+2}{2} \right) \left| f'\left(\tfrac{x_i+x_{i+1}}{2}\right) \right| + \left(\tfrac{s-2}{4} + 2\left(\tfrac{3}{4}\right)^{s+2} \right) |f'(x_{i+1})| \right). \end{aligned}$$

PROOF. Applying Theorem 2.9 on the subintervals $[x_i, x_{i+1}]$ $(i = 0,1,...,n-1)$ of the partition Υ, we get

$$\begin{aligned} & \left| \tfrac{1}{8}\left(f(x_i) + 6f\left(\tfrac{x_i+x_{i+1}}{2}\right) + f(x_{i+1})\right) - \tfrac{1}{x_{i+1}-x_i} \int_{x_i}^{x_{i+1}} f(t)\,dt \right| \\ \le \ & \tfrac{x_{i+1}-x_i}{4(s+1)(s+2)} \left(\left(\tfrac{s-2}{4} + 2\left(\tfrac{3}{4}\right)^{s+2} \right) |f'(x_i)| + \left(\left(\tfrac{1}{4}\right)^{s+1} + \tfrac{3s+2}{2} \right) \left| f'\left(\tfrac{x_i+x_{i+1}}{2}\right) \right| \right. \\ & \left. + \left(\tfrac{s-2}{4} + 2\left(\tfrac{3}{4}\right)^{s+2} \right) |f'(x_{i+1})| \right). \end{aligned} \tag{15}$$

Multiplying both sides of (15) by $(x_{i+1} - x_i)$, and then summing the obtained inequalities for all $i = 0,1,...,n-1$ and using the triangular inequality, we get the desired result.

Application to special means

For arbitrary real numbers a,b we have:

The Arithmetic mean: $A(a,b) = \frac{a+b}{2}$.

The Geometric mean: $G(a,b) = \sqrt{ab}$, $a,b > 0$.

The p-Logarithmic mean: $L_p(a,b) = \left(\frac{b^{p+1}-a^{p+1}}{(p+1)(b-a)}\right)^{\frac{1}{p}}$, $a,b > 0, a \ne b$ and $p \in \mathbb{R} \setminus \{-1,0\}$.

PROPOSITION 4.2. Let $a,b \in \mathbb{R}$ with $0 < a < b$, then we have

$$\left|A\left(a^2,b^2\right)+3A^2(a,b)-4L_2^2(a,b)\right| \leq \frac{5(b-a)^2}{8}.$$

PROOF. The assertion follows from Theorem 3.1, applied to the function $f(x) = x^2$.

REFERENCES

M. A. Ali, H. Kara, J. Tariboon, S. Asawasamrit, H. Budak and F. Hezenci, Some New Simpson's-Formula-Type Inequalities for Twice-Differentiable Convex Functions via Generalized Fractional Operators. Symmetry, 13 (2021), no. 12, 2249.

M. U. Awan, M. Z. Javed, M. Th. Rassias, M. A. Noor and K. I. Noor, Simpson type inequalities and applications. J. Anal. 29 (2021), no. 4, 1403–1419.

N. Azzouza and B. Meftah, Some weighted integral inequalities for differentiable $beta$-convex functions. J. Interdiscip. Math. 24 (2021), no. 5, 1-22.

H. Budak, F. Hezenci and H. Kara, On parameterized inequalities of Ostrowski and Simpson type for convex functions via generalized fractional integrals. Math. Methods Appl. Sci. 44 (2021), no. 17, 12522–12536.

H. Budak, F. Hezenci and H. Kara, On generalized Ostrowski, Simpson and trapezoidal type inequalities for co-ordinated convex functions via generalized fractional integrals. Adv. Difference Equ. 2021, Paper no. 312, 32 pp.

W. W. Breckner, Stetigkeitsaussagen für eine Klasse verallgemeinerter konvexer Funktionen in topologischen linearen Räumen. (German) Publ. Inst. Math. (Beograd) (N.S.) 23 (37) (1978), 13-20.

T. Chiheb, N. Boumaza and B. Meftah, Some new Simpson-like type inequalities via preqausiinvexity. Transylv. J. Math. Mech.12 (2020), no.1, 1-10.

F. Hezenci, H. Budak and H. Kara, New version of fractional Simpson type inequalities for twice differentiable functions. Adv. Difference Equ. 2021, Paper no. 460, 10 pp.

İ. İşcan, Hermite-Hadamard and Simpson-like type inequalities for differentiable harmonically convex functions. J. Math. 2014, Art. ID 346305, 10 pp.

A. Kashuri, B. Meftah and P.O. Mohammed, Some weighted Simpson type inequalities for differentiable s-convex functions and their applications. J. Frac. Calc. &Nonlinear Sys.1 (2021) no. 1, 75-94.

A. A. Kilbas, H. M. Srivastava and J. J. Trujillo, Theory and applications of fractional differential equations. North-Holland Mathematics Studies, 204. Elsevier Science B.V., Amsterdam, 2006.

U. S. Kirmaci, Refinements of Hermite-Hadamard type inequalities for s-convex functions with applications to special means. Univ. J. Math. App, 4 (2021), no.3, 114-124.

W. Liu, Some Simpson type inequalities for h-convex and (α,m)-convex functions. J. Comput. Anal. Appl. 16 (2014), no. 5, 1005–1012.

C. Y. Luo, T. S. Du, M. Kunt and Y. Zhang, Certain new bounds considering the weighted Simpson-like type inequality and applications. J. Inequal. Appl. 2018, Paper no. 332, 20 pp.

J. E. Pečarić, F. Proschan and Y. L. Tong, Convex functions, partial orderings, and statistical applications. Mathematics in Science and Engineering, 187. Academic Press, Inc., Boston, MA, 1992.

M. Z. Sarikaya, E. Set and E. Özdemir, On new inequalities of Simpson's type for convex functions, RGMIA Research Report Collection, 13 (2010), no.2, article 2.

Y. Shuang, Y. Wang and F. Qi, Integral inequalities of Simpson's type for (α,m)-convex functions. J. Nonlinear Sci. Appl. 9 (2016), no. 12, 6364–6370.

X. You, F. Hezenci, H. Budak and H. Kara, New Simpson type inequalities for twice differentiable functions via generalized fractional integrals. AIMS Math. 7 (2022), no. 3, 3959–3971.

N. Kamouche
Department of Mathematics, Faculty of Sciences,
University of Badji Mokhtar–Annaba, Algeria.
Email: kamouche_nesrine@hotmail.fr

S. Ghomrani
Ecole normale supérieure de Sétif Messaoud Zeggar,
Département des sciences. Sétif , Algeria.
Email: sarra.ghomrani@hotmail.fr

B. Meftah
Département des Mathématiques
Faculté des mathématiques, de l'informatique et des sciences de la matière
Université 8 mai 1945 Guelma, Algeria
Email: badrimeftah@yahoo.fr

A note on generalized subclasses of multivalent quasi-convex functions

G. SINGH AND G. SINGH

Abstract

This paper is concerned with certain generalized subclasses of multivalent quasi-convex functions defined with subordination. Various properties of these classes in regard to the coefficient estimates, distortion theorems, growth theorems, argument theorems and inclusion relations are discussed. Also, the relations with the earlier known results are established.

Mathematics Subject Classification 2010: 30C45, 30C50
General Terms: Regression, Generalized linear model, Model choice, Model selection, Information theory **Keywords:** Univalent functions, Subordination, multivalent functions, close-to-convex functions, quasi-convex functions.

1. INTRODUCTION

By $\mathbb{C}$, we denote the complex plane and the unit disc is defined as $E = \{z : z \in \mathbb{C}, |z| < 1\}$. $\mathscr{A}_p(p \geq 1)$, denotes the class of analytic functions f in the open unit disc E which has a Taylor series of the form

$$f(z) = z^p + \sum_{k=p+1}^{\infty} a_k z^k. \tag{1}$$

For $p = 1$, the class $\mathscr{A}_p$ agrees with $\mathscr{A}_1$, which is the class of analytic functions of the form $f(z) = z + \sum_{k=2}^{\infty} a_k z^k$ and normalized by the conditions $f(0) = f'(0) - 1 = 0$. The class of functions in $\mathscr{A}_1$ and which are univalent in E, is denoted by $\mathscr{S}$.

The class of Schwarzian functions is denoted by $\mathscr{U}$ and it consists of analytic functions of the form

$$w(z) = \sum_{k=1}^{\infty} c_k z^k,$$

which satisfy the conditions $w(0) = 0, |w(z)| < 1$ in the unit disc E. It was proved in [16] that for $w \in \mathscr{U}$, $|c_1| \leq 1$ and $|c_2| \leq 1 - |c_1|^2$.

10.2478/jamsi-2022-0007

sciendo

For two analytic functions f and g in E, f is said to be subordinate to g if a Schwarz function $w \in \mathscr{U}$ can be found, such that $f(z) = g(w(z))$ and it is denoted by $f \prec g$. Further, if the function g is univalent in E, then $f \prec g$ is equivalent to $f(0) = g(0)$ and $f(E) \subset g(E)$.

Goluzina [6] established the classes $\mathscr{S}_p^*(\alpha)$ and $\mathscr{K}_p(\alpha)$ $(0 \leq \alpha < p)$, which are the subclasses of $\mathscr{A}_p$ that are respectively the classes of p-valently starlike functions and p-valently convex functions of order α and defined as

$$\mathscr{S}_p^*(\alpha) = \left\{f : f \in \mathscr{A}_p, Re\left(\frac{zf'(z)}{f(z)}\right) > \alpha, z \in E\right\}$$

and

$$\mathscr{K}_p(\alpha) = \left\{f : f \in \mathscr{A}_p, Re\left(\frac{(zf'(z))'}{f'(z)}\right) > \alpha, z \in E\right\}.$$

For $0 \leq \alpha < 1$, $\mathscr{S}_1^*(\alpha) \equiv \mathscr{S}^*(\alpha)$ and $\mathscr{K}_1(\alpha) \equiv \mathscr{K}(\alpha)$, the classes of starlike functions of order α and convex functions of order α respectively, introduced by Robertson [21]. Also $\mathscr{S}_p^*(0) \equiv \mathscr{S}_p^*$ and $\mathscr{K}_p(0) \equiv \mathscr{K}_p$, the classes of p-valent starlike functions and p-valent convex functions, respectively. Further $\mathscr{S}_1^*(0) \equiv \mathscr{S}^*$ and $\mathscr{K}_1(0) \equiv \mathscr{K}$, the well known classes of starlike functions and convex functions, respectively.

$\mathscr{C}_p(\alpha)$ is the class of p-valent close-to-convex functions defined as

$$\mathscr{C}_p(\alpha) = \left\{f : f \in \mathscr{A}_p, Re\left(\frac{zf'(z)}{g(z)}\right) > \alpha, g \in \mathscr{S}_p^*, z \in E\right\}.$$

This class was established by Umezawa [27]. For $p = 1$, $\alpha = 0$, the class $\mathscr{C}_p(\alpha)$ reduces to $\mathscr{C}$, the class of close-to-convex functions introduced by Kaplan [9].

Noor [17] introduced the class $\mathscr{C}^*$ of quasi-convex functions. A function $f \in \mathscr{A}_1$ is said to be quasi-convex if there exists a convex function $h \in \mathscr{K}$ such that $Re\left(\frac{(zf'(z))'}{h'(z)}\right) > 0$. Every quasi-convex function is convex and so univalent. Different subclasses of quasi-convex functions were studied by various authors including Selvaraj and Stelin [23], Selvaraj et al. [24], Xiong and Liu [28] and Singh and Singh [26].

The corresponding class of p-valent quasi convex functions is defined as below:

$$\mathscr{C}_p^* = \left\{f : f \in \mathscr{A}_p, Re\left(\frac{(zf'(z))'}{h'(z)}\right) > 0, h \in \mathscr{K}_p, z \in E\right\}.$$

Obviously $\mathscr{C}_1^* \equiv \mathscr{C}^*$. Not much work has done related to the study of subclasses of multivalent quasi-convex functions.

For $-1 \leq B < A \leq 1$ and $0 \leq \alpha < p$, Aouf [2] introduced the class $\mathscr{P}(A,B;p;\alpha)$, which is a subclass of $\mathscr{A}_p$ consisting of the functions of the form $p(z) = p + \sum_{k=1}^{\infty} p_k z^k$ such that $p(z) \prec \frac{p+[pB+(A-B)(p-\alpha)]z}{1+Bz}$. For $p = 1$, $\mathscr{P}(A,B;p;\alpha)$ reduces to $\mathscr{P}(A,B;\alpha)$, the class introduced by Polatoglu et al. [19]. Also for $p = 1, \alpha = 0$, the class $\mathscr{P}(A,B;p;\alpha)$ agrees with $\mathscr{P}(A,B)$, which is a subclass of $\mathscr{A}_1$ introduced by Janowski [8].

Again, for $-1 \leq B < A \leq 1$ and $0 \leq \alpha < p$, Aouf [2; 4], introduced the following useful classes:

$$\mathscr{S}^*(A,B;p;\alpha) = \left\{ f : f \in \mathscr{A}_p, \frac{zf'(z)}{f(z)} \prec \frac{p+[pB+(A-B)(p-\alpha)]z}{1+Bz}, z \in E \right\}$$

and

$$\mathscr{K}(A,B;p;\alpha) = \left\{ f : f \in \mathscr{A}_p, \frac{(zf'(z))'}{f'(z)} \prec \frac{p+[pB+(A-B)(p-\alpha)]z}{1+Bz}, z \in E \right\}.$$

The following points are to be noted:

(i) $\mathscr{S}^*(1,-1;p;\alpha) \equiv \mathscr{S}_p^*(\alpha)$ and $\mathscr{K}(1,-1;p;\alpha) \equiv \mathscr{K}_p(\alpha)$.

(ii) $\mathscr{S}^*(A,B;p;0) \equiv \mathscr{S}_p^*(A,B)$ and $\mathscr{K}(A,B;p;0) \equiv \mathscr{K}_p(A,B)$, the classes studied by Hayami and Owa [7].

(iii) $\mathscr{S}^*(A,B;1;\alpha) \equiv \mathscr{S}^*(A,B;\alpha)$, the class studied by Polatoglu et al. [19].

(iv) $\mathscr{S}^*(A,B;1;0) \equiv \mathscr{S}^*(A,B)$ and $\mathscr{K}(A,B;1;0) \equiv \mathscr{K}(A,B)$, the subclasses of starlike and convex functions respectively, introduced by Janowski [8] and studied further by Goel and Mehrok [5].

(v) $\mathscr{S}^*(1,-1;1;\alpha) \equiv \mathscr{S}^*(\alpha)$ and $\mathscr{K}(1,-1;1;\alpha) \equiv \mathscr{K}(\alpha)$.

(vi) $\mathscr{S}^*(1,-1;1;0) \equiv \mathscr{S}^*$ and $\mathscr{K}(1,-1;1;0) \equiv \mathscr{K}$.

Throughout this paper, we assume that $-1 \leq D < C \leq 1$, $-1 \leq B < A \leq 1$, $0 \leq \alpha < p$, $0 \leq \beta < p$ and $z \in E$.

Getting motivated by the above mentioned work, now we are on the stage to define the following classes:

DEFINITION 1. Let $\mathscr{C}^*(A,B;C,D;p;\beta;\alpha)$ denote the class of functions $f \in \mathscr{A}_p$ and satisfying the condition

$$\frac{(zf'(z))'}{h'(z)} \prec \frac{p+[pD+(C-D)(p-\beta)]z}{1+Dz},$$

where

$$h(z) = z^p + \sum_{k=p+1}^{\infty} b_k z^k \in \mathscr{K}(A,B;p;\alpha).$$

DEFINITION 2. $\mathscr{C}_s^*(A,B;C,D;p;\beta;\alpha)$ is the class of functions $f \in \mathscr{A}_p$ which satisfy the condition

$$\frac{(zf'(z))'}{g'(z)} \prec \frac{p+[pD+(C-D)(p-\beta)]z}{1+Dz},$$

where

$$g(z) = z^p + \sum_{k=p+1}^{\infty} d_k z^k \in \mathscr{S}^*(A,B;p;\alpha).$$

The following observations are obvious:

(i) $\mathscr{C}^*(A,B;C,D;p;0;0) \equiv \mathscr{C}^*(A,B;C,D;p)$.

(ii) $\mathscr{C}^*(A,B;C,D;1;0;0) \equiv \mathscr{C}^*(A,B;C,D)$, the subclass of quasi-convex functions investigated by Singh and Singh [26].

(iii) $\mathscr{C}^*(1,-1;C,D;1;0;0) \equiv \mathscr{C}^*(C,D)$, the class studied by Xiong and Liu [28].

(iv) $\mathscr{C}^*(1,-1;1,(1-2\alpha)\beta;\beta;1;0;0) \equiv \mathscr{C}^*(\alpha,\beta)$, the subclass of quasi-convex functions introduced by Selvaraj and Stelin [23].

(v) $\mathscr{C}^*(1,-1;1,-1;1;0;0) \equiv \mathscr{C}^*$.

(vi) $\mathscr{C}_s^*(A,B;C,D;p;0;0) \equiv \mathscr{C}_s^*(A,B;C,D;p)$.

(vii) $\mathscr{C}_s^*(A,B;C,D;1;0;0) \equiv \mathscr{C}_s^*(A,B;C,D)$, the subclass of quasi-convex functions investigated by Singh and Singh [26].

(viii) $\mathscr{C}_s^*(1,-1;C,D;1;0;0) \equiv \mathscr{C}_s^*(C,D)$, the class discussed by Singh and Singh [26].

(ix) $\mathscr{C}_s^*(1,-1;1,(1-2\alpha)\beta;\beta;1;0;0) \equiv \mathscr{C}_s^*(\alpha,\beta)$, the subclass of quasi-convex functions studied by Selvaraj et al. [24].

(x) $\mathscr{C}_s^*(1,-1;1,-1;1;0;0) \equiv \mathscr{C}_s^*$, the subclass of quasi-convex functions discussed by Singh and Singh [26].

Obsah...

LEMMA 1. [2] If $P(z) = \frac{p+[pD+(C-D)(p-\beta)]w(z)}{1+Dw(z)} = p + \sum_{k=1}^{\infty} p_k z^k \in \mathscr{P}(C,D;p;\beta)$, then

$$|p_n| \leq (C-D)(p-\beta), n \geq p.$$

The bounds are sharp for $w(z) = z^n$ and for the function

$$P(z) = p + (C-D)(p-\beta)z^n + ...$$

LEMMA 2. [13] Let $-1 \le D_2 \le D_1 < C_1 \le C_2 \le 1$, then

$$\tfrac{1+C_1z}{1+D_1z} \prec \tfrac{1+C_2z}{1+D_2z}.$$

LEMMA 3. [22] If $\psi(z)$ is regular in E, $\phi(z)$ and $h(z)$ are convex univalent in E such that $\psi(z) \prec \phi(z)$, then $\psi(z) * h(z) \prec \phi(z) * h(z)$, $z \in E$.

In this paper, we investigate various properties such as the coefficient estimates, distortion theorems, growth theorems, argument theorems and inclusion relations for the classes $\mathscr{C}^*(A,B;C,D;p;\beta;\alpha)$ and $\mathscr{C}_s^*(A,B;C,D;p;\beta;\alpha)$. The results already proved by various authors will follow as special cases.

2. STUDY OF THE CLASS $\mathscr{C}^*(A,B;C,D;P;\beta;\alpha)$

THEOREM 1. Let $f \in \mathscr{C}^*(A,B;C,D;p;\beta;\alpha)$, then for $n \ge p+1$,

$$|a_n| \le \tfrac{p^2}{n^2[(n-p)!]}\Pi_{k=0}^{n-(p+1)}|(B-A)(p-\alpha)+Bk|$$
$$+\frac{(C-D)(p-\beta)}{n^2}\left[p+\sum_{m=p+1}^{n-1}\frac{p}{(m-p)!}\Pi_{k=0}^{m-(p+1)}|(B-A)(p-\alpha)+Bk|\right]. \tag{2}$$

The result is sharp.

PROOF. For $f \in \mathscr{C}^*(A,B;C,D;p;\beta;\alpha)$, we have

$$(zf'(z))' = h'(z)P(z), \tag{3}$$

where

$$h(z) = z^p + \textstyle\sum_{k=p+1}^{\infty} b_k z^k \in \mathscr{K}(A,B;p;\alpha)$$

and

$$P(z) = p + \textstyle\sum_{k=1}^{\infty} p_k z^k \in \mathscr{P}(C,D;p;\beta).$$

Expansion of (3), yields

$$p^2+(p+1)^2a_{p+1}z+(p+2)^2a_{p+2}z^2+\ldots+n^2a_nz^{n-p}+\ldots$$
$$=[p+(p+1)b_{p+1}z+(p+2)b_{p+2}z^2+\ldots+nb_nz^{n-p}+\ldots][p+p_1z+p_2z^2+\ldots+p_{n-1}z^{n-1}+\ldots]. \tag{4}$$

On equating the coefficients of z^{n-p} in (4), we have

$$n^2a_n = pnb_n+(n-1)p_1b_{n-1}+(n-2)p_2b_{n-2}+\ldots+2p_{n-2}b_2+pp_{n-p}. \tag{5}$$

Application of triangle inequality and using Lemma 1 in (5), it gives

$$n^2|a_n| \le pn|b_n| + (C-D)(p-\beta)\left[(n-1)|b_{n-1}| + (n-2)|b_{n-2}| + \dots + (p+1)|b_{p+1}| + p\right]. \quad (6)$$

It was proved in [4] that, for $h(z) = z^p + \sum_{k=p+1}^{\infty} b_k z^k \in \mathscr{K}(A,B;p;\alpha)$,

$$|b_n| \le \frac{p}{n[(n-p)!]} \Pi_{j=0}^{n-(p+1)} |(B-A)(p-\alpha) + Bj|, n \ge p+1. \quad (7)$$

Using (7) in (6), the result (2) can be easily derived.
Equality sign in (2) is attained for the functions $f_n(z)$ defined by

$$(zf_n'(z))' = pz^{p-1}(1-B\delta_1 z)^{\frac{(A-B)(p-\alpha)}{B}} \left[\frac{p + \{pD + (C-D)(p-\beta)\}\delta_2 z}{1+D\delta_2 z}\right], |\delta_1| = |\delta_2| = 1. \quad (8)$$

REMARK 1. (i) On putting $\alpha = 0, \beta = 0$ in Theorem 1, the result for the class $\mathscr{C}^*(A,B;C,D;p)$ can be easily obtained.
(ii) For $\alpha = 0, \beta = 0, p = 1$, Theorem 1 gives the result established by Singh and Singh [26].
(iii) By giving the values $A = 1, B = -1, \alpha = 0, \beta = 0, p = 1$, the result due to Xiong and Liu [28], can be easily obtained from Theorem 1.
(iv) Substituting for $A = 1, B = -1, C = (1-2\alpha)\beta, D = \beta, \alpha = 0, \beta = 0, p = 1$ in Theorem 1, we can easily get the result due to Selvaraj and Stelin [23].
(v) For $A = 1, B = -1, C = 1, D = -1, \alpha = 0, \beta = 0, p = 1$, the result established by Noor [17], can be easily obtained from Theorem 1.

THEOREM 2. If $f \in \mathscr{C}^*(A,B;C,D;p;\beta;\alpha)$, then for $|z| = r, 0 < r < 1$, we have for $B \neq 0$,

$$\frac{1}{r}\int_0^r pt^{p-1}(1-Bt)^{\frac{(A-B)(p-\alpha)}{B}} \left[\frac{p - \{pD + (C-D)(p-\beta)\}t}{1-Dt}\right] dt \le |f'(z)|$$

$$\le \frac{1}{r}\int_0^r pt^{p-1}(1+Bt)^{\frac{(A-B)(p-\alpha)}{B}} \left[\frac{p + \{pD + (C-D)(p-\beta)\}t}{1+Dt}\right] dt; \quad (9)$$

$$\int_0^r \left[\frac{1}{s}\int_0^s pt^{p-1}(1-Bt)^{\frac{(A-B)(p-\alpha)}{B}} \left[\frac{p - \{pD + (C-D)(p-\beta)\}t}{1-Dt}\right] dt\right] ds \le |f(z)|$$

$$\le \int_0^r \left[\frac{1}{s}\int_0^s pt^{p-1}(1+Bt)^{\frac{(A-B)(p-\alpha)}{B}} \left[\frac{p + \{pD + (C-D)(p-\beta)\}t}{1+Dt}\right] dt\right] ds, \quad (10)$$

for $B = 0$,

$$\frac{1}{r}\int_0^r pt^{p-1}e^{-A(p-\alpha)t} \left[\frac{p - \{pD + (C-D)(p-\beta)\}t}{1-Dt}\right] dt \le |f'(z)|$$

$$\le \frac{1}{r}\int_0^r pt^{p-1}e^{A(p-\alpha)t} \left[\frac{p + \{pD + (C-D)(p-\beta)\}t}{1+Dt}\right] dt; \quad (11)$$

$$\int_0^r \left[\frac{1}{s}\int_0^s pt^{p-1}e^{-A(p-\alpha)t}\left[\frac{p-\{pD+(C-D)(p-\beta)\}t}{1-Dt}\right]dt\right]ds \le |f(z)|$$

$$\le \int_0^r \left[\frac{1}{s}\int_0^s pt^{p-1}e^{A(p-\alpha)t}\left[\frac{p+\{pD+(C-D)(p-\beta)\}t}{1+Dt}\right]dt\right]ds. \quad (12)$$

Estimates are sharp.

PROOF. Equation (3) can be expressed as

$$|(zf'(z))'| = |h'(z)||P(z)|. \quad (13)$$

It was established in [4] that, for $P(z) \in \mathscr{P}(A,B;p;\alpha)$,

$$\frac{p-[pD+(C-D)(p-\beta)]r}{1-Dr} \le |P(z)| \le \frac{p+[pD+(C-D)(p-\beta)]r}{1+Dr}. \quad (14)$$

Aouf [4] proved that, for $h(z) \in \mathscr{K}(A,B;p;\alpha)$,

$$pr^{p-1}(1-Br)^{\frac{A-B}{B}(p-\alpha)} \le |h'(z)| \le pr^{p-1}(1+Br)^{\frac{A-B}{B}(p-\alpha)} \text{ if } B \neq 0,$$

and

$$pr^{p-1}e^{-A(p-\alpha)r} \le |h'(z)| \le pr^{p-1}e^{A(p-\alpha)r} \text{ if } B = 0.$$

Using the above inequalities and (14) in (13), the results (9) and (11) can be easily obtained. On integrating (9) and (11) from 0 to r, the results (10) and (12) are obvious. Sharpness follows for the functions $f_n(z)$ defined as

$$(zf_n'(z))' = pz^{p-1}(1+B\delta_3 z)^{\frac{(A-B)(p-\alpha)}{B}}\left[\frac{p+\{pD+(C-D)(p-\beta)\}\delta_4 z}{1+D\delta_4 z}\right] \text{ if } B \neq 0,$$

$$(zf_n'(z))' = pz^{p-1}e^{A(p-\alpha)\delta_5 z}\left[\frac{p+\{pD+(C-D)(p-\beta)\}\delta_4 z}{1+D\delta_4 z}\right] \text{ if } B = 0,$$

where $|\delta_3| = |\delta_4| = |\delta_5| = 1$.

REMARK 2. (i) Putting $\alpha = 0, \beta = 0$ in Theorem 2, the result for the class $\mathscr{C}^*(A,B;C,D;p)$ can be easily obtained.
(ii) For $\alpha = 0, \beta = 0, p = 1$, Theorem 2 gives the result established by Singh and Singh [26].
(iii) By giving the values $A = 1, B = -1, \alpha = 0, \beta = 0, p = 1$, the result due to Xiong and Liu [28], can be easily obtained from Theorem 2.
(iv) Substituting for $A = 1, B = -1, C = (1-2\alpha)\beta, D = \beta, \alpha = 0, \beta = 0, p = 1$ in Theorem 2, we can easily get the result due to Selvaraj and Stelin [23].
(v) For $A = 1, B = -1, C = 1, D = -1, \alpha = 0, \beta = 0, p = 1$, the result established by Noor [17], can be easily obtained from Theorem 2.

THEOREM 3. If $f \in \mathscr{C}^*(A,B;C,D;p;\beta;\alpha)$ and let $F(z) = zf'(z)$, then

$$\left|arg\tfrac{F'(z)}{z^{p-1}}\right| \leq \tfrac{(A-B)(p-\alpha)}{B} sin^{-1}(Br) + sin^{-1}\left(\tfrac{(C-D)(p-\beta)r}{p-[pD+(C-D)(p-\beta)]Dr^2}\right) \text{ if } B \neq 0,$$

and

$$\left|arg\tfrac{F'(z)}{z^{p-1}}\right| \leq A(p-\alpha)r + sin^{-1}\left(\tfrac{(C-D)(p-\beta)r}{p-[pD+(C-D)(p-\beta)]Dr^2}\right) \text{ if } B = 0.$$

The bounds are sharp.

PROOF. (3) can be expressed as

$$F'(z) = h'(z)P(z).$$

Therefore, we have

$$\left|arg\frac{F'(z)}{z^{p-1}}\right| \leq |argP(z)| + \left|arg\frac{pf_1(z)}{z^p}\right|, \tag{15}$$

where $f_1(z) = \frac{zh'(z)}{p}$.
Aouf [2], established that for $P(z) \in \mathscr{P}(A,B;p;\alpha)$,

$$|argP(z)| \leq sin^{-1}\left(\frac{(C-D)(p-\beta)r}{p-[pD+(C-D)(p-\beta)]Dr^2}\right). \tag{16}$$

It was proved by Aouf [4], that

$$\left|arg\tfrac{pf_1(z)}{z^p}\right| \leq \tfrac{(A-B)(p-\alpha)}{B} sin^{-1}(Br) \text{ if } B \neq 0,$$

$$\left|arg\tfrac{pf_1(z)}{z^p}\right| \leq A(p-\alpha)r \text{ if } B = 0.$$

Using the above inequalities and (16) in (15), the proof of Theorem 3 is obvious. Results are sharp for the extremal function defined in Theorem 1.

REMARK 3. (i) On putting $\alpha = 0, \beta = 0$ in Theorem 3, the result for the class $\mathscr{C}^*(A,B;C,D;p)$ can be easily obtained.
(ii) By giving the values $A = 1, B = -1, \alpha = 0, \beta = 0, p = 1$, the result due to Xiong and Liu [28] can be easily obtained from Theorem 3.
(iii) Substituting for $A = 1, B = -1, C = (1-2\alpha)\beta, D = \beta, \alpha = 0, \beta = 0, p = 1$ in Theorem 3, we can easily get the result due to Selvaraj and Stelin [23].
(iv) For $A = 1, B = -1, C = 1, D = -1, \alpha = 0, \beta = 0, p = 1$, the result established by Noor [17], can be easily obtained from Theorem 3.

THEOREM 4. If $f \in \mathscr{C}^*(A,B;C,D;p;\beta;\alpha)$, then

$$|a_{p+1}| \leq \tfrac{p}{(p+1)^2}[(p-\beta)(C-D)+(p-\alpha)(A-B)\{p+(C-D)(p-\beta)\}],$$

and

$$|a_{p+2}| \leq \tfrac{p}{(p+2)^2}\Big[(p-\beta)(C-D)+(A-B)(p-\alpha)\{[p+(C-D)(p-\beta)]$$
$$+|(C-D)(p-\beta)(1-D)|\}\Big] \text{ if } |(A-B)(p-\alpha)-B| \leq 1,$$

and

$$|a_{p+2}| \leq \tfrac{p}{(p+2)^2}\Big[(p-\beta)(C-D)+(A-B)(p-\alpha)\{[p+(C-D)(p-\beta)]$$
$$+|(C-D)(p-\beta)(1-D)|\}+\tfrac{[p+(C-D)(p-\beta)]+|(A-B)(p-\alpha)-B|}{2}\Big] \text{ if } |(A-B)(p-\alpha)-B| > 1.$$ The estimates are sharp.

PROOF. Using the principle of subordination in Definition 1, we obtain

$$\tfrac{(zf'(z))'}{h'(z)} = \tfrac{p+[pD+(C-D)(p-\beta)]w(z)}{1+Dw(z)}.$$

On expanding and comparing the coefficients, it leads to

$$a_{p+1} = \frac{p+(C-D)(p-\beta)}{p+1}b_{p+1} + \frac{p(C-D)(p-\beta)}{(p+1)^2}c_1 \tag{17}$$

and

$$a_{p+2} = \tfrac{p+(C-D)(p-\beta)}{p+2}b_{p+2}$$
$$+\frac{(p+1)}{(p+2)^2}[(C-D)(p-\beta)(1-D)]b_{p+1}c_1 + \frac{p(C-D)(p-\beta)}{(p+2)^2}[c_2-Dc_1^2]. \tag{18}$$

Aouf [4] proved that for $h(z) = z^p + \sum_{k=p+1}^{\infty} b_k z^k \in \mathscr{K}(A,B;p;\alpha)$,

$$|b_{p+1}| \leq \frac{p(A-B)(p-\alpha)}{p+1} \tag{19}$$

and

$$|b_{p+2}| \leq \frac{p(A-B)(p-\alpha)}{2(p+2)} max\{1, |(A-B)(p-\alpha)p-B|\}. \tag{20}$$

Also it was proved in [10], that for any complex number γ,

$$|c_2-\gamma c_1^2| \leq max\{1,|\gamma|\}. \tag{21}$$

Using (19), (20) and (21) along with the inequality $|c_1| \leq 1$ in (17) and (18), the results are obvious.
The results are sharp for the extremal functions defined in (8).

THEOREM 5. Let $-1 \le D_2 = D_1 < C_1 \le C_2 \le 1$ and $0 \le \beta_2 \le \beta_1 < p$, then

$$\mathscr{C}^*(A,B;C_1,D_1;p;\beta_1;\alpha) \subset \mathscr{C}^*(A,B;C_2,D_2;p;\beta_2;\alpha).$$

PROOF. As $f \in \mathscr{C}^*(A,B;C_1,D_1;p;\beta_1;\alpha)$, so

$$\frac{(zf'(z))'}{h'(z)} \prec \frac{p+[pD_1+(C_1-D_1)(p-\beta_1)]z}{1+D_1z}.$$

As $-1 \le D_2 = D_1 < C_1 \le C_2 \le 1$ and $0 \le \beta_2 \le \beta_1 < p$, we have

$$-1 \le D_1 + \frac{(p-\beta_1)(C_1-D_1)}{p} \le D_2 + \frac{(p-\beta_2)(C_2-D_2)}{p} \le 1.$$

So by Lemma 2, we obtain

$$\frac{(zf'(z))'}{h'(z)} \prec \frac{p+[pD_2+(C_2-D_2)(p-\beta_2)]z}{1+D_2z},$$

which implies $f \in \mathscr{C}^*(A,B;C_2,D_2;p;\beta_2;\alpha)$.

THEOREM 6. If $f \in \mathscr{C}^*(A,B;C,D;p;\beta;\alpha)$, then there exists $P(z) \in \mathscr{P}(C,D;p;\alpha)$ such that for all s and t with $|s| \le 1$, $|t| \le 1$ $(s \ne t)$,

$$\frac{(szf'(sz))'P(tz)(tz)^{p-1}}{(tzf'(tz))'P(sz)(sz)^{p-1}} = \left(\frac{1+Bsz}{1+Btz}\right)^{\left(\frac{A-B}{B}\right)(p-\alpha)}, \text{ if } B \ne 0,$$

and

$$\frac{(szf'(sz))'P(tz)(tz)^{p-1}}{(tzf'(tz))'P(sz)(sz)^{p-1}} = e^{A(p-\alpha)(s-t)z} \text{ if } B = 0.$$

PROOF. Firstly assume that $B \ne 0$.
From definition, we have

$$(zf'(z))' = P(z)h'(z).$$

On differentiating logarithmically, it yields

$$\frac{z(zf'(z))''}{(zf'(z))'} - \frac{zP'(z)}{P(z)} - p + 1 = 1 + \frac{zh''(z)}{h'(z)} - p.$$

As $h \in K(A,B;p;\alpha)$, therefore

$$\frac{z(zf'(z))''}{(zf'(z))'} - \frac{zP'(z)}{P(z)} - p + 1 \prec \frac{(A-B)(p-\alpha)z}{1+Bz},$$

where $\frac{(A-B)(p-\alpha)z}{1+Bz}$ is convex, univalent in E. For $|s| \le 1$, $|t| \le 1$ $(s \ne t)$,

$$h(z) = \int_0^z \left(\frac{s}{1-su} - \frac{t}{1-tu}\right) du$$

is convex univalent in E. Using Lemma 3, we have

$$\left(\frac{z(zf'(z))''}{(zf'(z))'} - \frac{zP'(z)}{P(z)} - p + 1\right) * h(z) \prec \frac{(A-B)(p-\alpha)z}{1+Bz} * h(z).$$

For any function $q(z)$ analytic in E with $q(0)=0$, we obtain

$$(q*h)(z)=\int_{tz}^{sz}q(u)\frac{du}{u}, z\in E.$$

Therefore, we have

$$\int_{tz}^{sz}\left(\frac{u(uf'(u))''}{(uf'(u))'}-\frac{uP'(u)}{P(u)}-p+1\right)\frac{du}{u}\prec(A-B)(p-\alpha)\int_{tz}^{sz}\frac{du}{1+Bu},$$

which follows the result. On the same lines, we can easily prove the result for $B=0$.

3. STUDY OF THE CLASS $\mathscr{C}_S^*(A,B;C,D;P;\beta;\alpha)$

THEOREM 7. Let $f(z)\in\mathscr{C}_s^*(A,B;C,D;p;\beta;\alpha)$, then for $n\geq p+1$,

$$|a_n|\leq\frac{p}{n}\Pi_{k=0}^{n-(p+1)}\frac{|(B-A)(p-\alpha)+Bk|}{k+1}$$
$$+\frac{(C-D)(p-\beta)}{n^2}\left[p+\sum_{m=p+1}^{n-1}m\Pi_{k=0}^{m-(p+1)}\frac{|(B-A)(p-\alpha)+Bk|}{k+1}\right]. \tag{22}$$

The result is sharp.

PROOF. Using the result due to Aouf [2] that, for $g(z)=z^p+\sum_{k=p+1}^{\infty}d_kz^k\in\mathscr{S}^*(A,B;p;\alpha)$,

$$|d_n|\leq\Pi_{j=0}^{n-(p+1)}\frac{|(B-A)(p-\alpha)+Bj|}{j+1}, n\geq p+1,$$

and following the procedure of Theorem 1, the proof is obvious.

Equality sign in (22) hold for the functions $f_n(z)$ defined by

$$(zf_n'(z))'$$
$$=z^{p-1}(1-B\delta_6z)^{\frac{(A-B)(p-\alpha)}{B}}\left[p-\frac{\delta_7z(A-B)(p-\alpha)}{1-B\delta_7z}\right]\left[\frac{p+\{pD+(C-D)(p-\beta)\}\delta_8z}{1+D\delta_8z}\right], \tag{23}$$

where $|\delta_6|=|\delta_7|=|\delta_8|=1$.

REMARK 4. (i) For $\alpha=0,\beta=0$, Theorem 7 gives the result for the class $\mathscr{C}_s^*(A,B;C,D;p)$.
(ii) For $\alpha=0,\beta=0,p=1$, Theorem 7 yields the result due to Singh and Singh [26].
(iii) Putting $A=1,B=-1,\alpha=0,\beta=0,p=1$ in Theorem 7, it yields the result for the class $\mathscr{C}_s^*(C,D)$.
(iv) Substituting for $A=1,B=-1,C=(1-2\alpha)\beta,D=\beta,\alpha=0,\beta=0,p=1$, the result due to Selvaraj et al. [24], can be easily obtained from Theorem 7.
(v) For $A=1,B=-1,C=1,D=-1,\alpha=0,\beta=0,p=1$, Theorem 7 gives the result for the class $\mathscr{C}_s^*$.

THEOREM 8. If $f \in \mathscr{C}_s^*(A,B;C,D;p;\beta;\alpha)$, then for $|z| = r, 0 < r < 1$, we have for $B \neq 0$,

$$\frac{1}{r}\int_0^r t^p(1-Bt)^{\frac{A-B}{B}(p-\alpha)}\left[\frac{p}{t}-\frac{(A-B)(p-\alpha)}{1-Bt}\right]\left[\frac{p-\{pD+(C-D)(p-\beta)\}t}{1-Dt}\right]dt \leq |f'(z)|$$

$$\leq \frac{1}{r}\int_0^r t^p(1+Bt)^{\frac{A-B}{B}(p-\alpha)}\left[\frac{p}{t}+\frac{(A-B)(p-\alpha)}{1+Bt}\right]\left[\frac{p+\{pD+(C-D)(p-\beta)\}t}{1+Dt}\right]dt; \quad (24)$$

$$\int_0^r\left[\frac{1}{s}\int_0^s t^p(1-Bt)^{\frac{(A-B)(p-\alpha)}{B}}\left[\frac{p}{t}-\frac{(A-B)(p-\alpha)}{1-Bt}\right]\left[\frac{p-\{pD+(C-D)(p-\beta)\}t}{1-Dt}\right]dt\right]ds \leq |f(z)|$$

$$\leq \int_0^r\left[\frac{1}{s}\int_0^r t^p(1+Bt)^{\frac{(A-B)(p-\alpha)}{B}}\left[\frac{p}{t}+\frac{(A-B)(p-\alpha)}{1+Bt}\right]\left[\frac{p+\{pD+(C-D)(p-\beta)\}t}{1+Dt}\right]dt\right]ds, \quad (25)$$

for $B = 0$,

$$\frac{1}{r}\int_0^r t^{p-1}e^{-A(p-\alpha)t}[p-A(p-\alpha)t]\left[\frac{p-\{pD+(C-D)(p-\beta)\}t}{1-Dt}\right]dt \leq |f'(z)|$$

$$\leq \frac{1}{r}\int_0^r t^{p-1}e^{A(p-\alpha)t}[p+A(p-\alpha)t]\left[\frac{p+\{pD+(C-D)(p-\beta)\}t}{1+Dt}\right]dt; \quad (26)$$

$$\int_0^r\left[\frac{1}{s}\int_0^s t^{p-1}e^{-A(p-\alpha)t}[p-A(p-\alpha)t]\left[\frac{p-\{pD+(C-D)(p-\beta)\}t}{1-Dt}\right]dt\right]ds \leq |f(z)|$$

$$\leq \int_0^r\left[\frac{1}{s}\int_0^s t^{p-1}e^{A(p-\alpha)t}[p+A(p-\alpha)t]\left[\frac{p+\{pD+(C-D)(p-\beta)\}t}{1+Dt}\right]dt\right]ds. \quad (27)$$

Estimates are sharp.

PROOF. Following the procedure of Theorem 2 and using the result that, for $g \in \mathscr{S}^*(A,B;p;\alpha)$,

$$r^p(1-Br)^{\frac{(A-B)(p-\alpha)}{B}}\left[\frac{p}{r}-\frac{(A-B)(p-\alpha)}{1-Br}\right] \leq |g'(z)| \leq r^p(1+Br)^{\frac{(A-B)(p-\alpha)}{B}}\left[\frac{p}{r}+\frac{(A-B)(p-\alpha)}{1+Br}\right] \text{ if } B \neq 0,$$

$$r^{p-1}e^{-A(p-\alpha)r}[p-A(p-\alpha)r] \leq |g'(z)| \leq r^{p-1}e^{A(p-\alpha)r}[p+A(p-\alpha)r] \text{ if } B = 0,$$

the results (24), (25), (26) and (27) can be easily derived.

Sharpness follows if we take $f_n(z)$ defined as

$$(zf_n'(z))' = z^p(1+B\delta_9 z)^{\frac{(A-B)(p-\alpha)}{B}}\left[\frac{p}{z}+\frac{(A-B)(p-\alpha)}{1+B\delta_{10}z}\right]\left[\frac{p+\{pD+(C-D)(p-\beta)\}\delta_{11}z}{1+D\delta_{11}z}\right] \text{ if } B \neq 0,$$

$$(zf_n'(z))' = z^{p-1}e^{A(p-\alpha)\delta_{12}z}\left[\frac{p}{z}+\frac{(A-B)(p-\alpha)}{1+B\delta_{10}z}\right]\left[\frac{p+\{pD+(C-D)(p-\beta)\}\delta_{11}z}{1+D\delta_{11}z}\right] \text{ if } B = 0,$$

where $|\delta_9| = |\delta_{10}| = |\delta_{11}| = |\delta_{12}| = 1$.

REMARK 5. (i) For $\alpha = 0, \beta = 0$, Theorem 8 gives the result for the class $\mathscr{C}_s^*(A,B;C,D;p)$.

(ii) For $\alpha = 0, \beta = 0, p = 1$, Theorem 8 yields the result due to Singh and Singh [26].

(iii) Putting $A = 1, B = -1, \alpha = 0, \beta = 0, p = 1$ in Theorem 8, it yields the result for the class $\mathscr{C}_s^*(C,D)$.

(iv) Substituting for $A=1, B=-1, C=(1-2\alpha)\beta, D=\beta, \alpha=0, \beta=0, p=1$, the result due to Selvaraj et al. [24], can be easily obtained from Theorem 8.
(v) For $A=1, B=-1, C=1, D=-1, \alpha=0, \beta=0, p=1$, Theorem 8 gives the result for the class $\mathscr{C}_s^*$.

THEOREM 9. If $f \in \mathscr{C}_s^*(A,B;C,D;p;\beta;\alpha)$, then

$$|a_{p+1}| \leq \frac{1}{(p+1)^2}[p(p-\beta)(C-D)+(p-\alpha)(A-B)\{p+(C-D)(p-\beta)\}], \tag{28}$$

and
$|a_{p+2}| \leq \frac{(A-B)(p-\alpha)}{p+2}\left[\frac{p(p+2)+(p-\beta)(C-D)}{2}+\frac{(C-D)(p-\beta)(1-D)}{p+2}\right]$
$+\frac{p(C-D)(p-\beta)}{(p+2)^2}$ if $|(A-B)(p-\alpha)-B| \leq p+1$, and
$|a_{p+2}| \leq \frac{(A-B)(p-\alpha)}{p+2}\left[\frac{[p(p+2)+(p-\beta)(C-D)][|(A-B)(p-\alpha)-B|]}{2(p+1)}+\frac{(C-D)(p-\beta)(1-D)}{p+2}\right]$
$+\frac{p(C-D)(p-\beta)}{(p+2)^2}$ if $|(A-B)(p-\alpha)-B| > p+1$. The bounds are sharp.

PROOF. For $g \in \mathscr{S}^*(A,B;p;\alpha)$,

$$|d_{p+1}| \leq \frac{(A-B)(p-\alpha)}{p+1} \tag{29}$$

and

$$|d_{p+2}| \leq \frac{(A-B)(p-\alpha)}{2} max\left\{1, \frac{|(A-B)(p-\alpha)-B|}{p+1}\right\}. \tag{30}$$

Using (29) and (30) and following the procedure of Theorem 4, the proof is obvious. The bounds are sharp for the function defined in (23).

THEOREM 10. Let $-1 \leq D_2 = D_1 < C_1 \leq C_2 \leq 1$ and $0 \leq \beta_2 \leq \beta_1 < p$, then

$$\mathscr{C}_s^*(A,B;C_1,D_1;p;\beta_1;\alpha) \subset \mathscr{C}_s^*(A,B;C_2,D_2;p;\beta_2;\alpha).$$

PROOF. Following the procedure of Theorem 5 and using Lemma 2, the proof is obvious.

REFERENCES

H. R. Abdel-Gawad and D. K. Thomas, A subclass of close-to-convex functions, *Publ. Inst. Math. (N. S.)*, 49(63)(1991), 61-66.

M. K. Aouf, On a class of p-valent starlike functions of order α, *Int. J. Math. Math. Sci.*, 10(4)(1987), 733-744.

M. K. Aouf, On a class of p-valent close-to-convex functions of order β and type α, *Int. J. Math. Math. Sci.*, 11(2)(1988), 259-266.

M. K. Aouf, On subclasses of p-valent close-to-convex functions, *Tamkang J. Math.*, 22(2)(1991), 133-143.

R. M. Goel and B. S. Mehrok, On a class of close-to-convex functions, *Ind. J. Pure Appl. Math.*, 12(5)(1981), 648-658.

E. G. Goluzina, On the coefficients of a class of functions, regular in a disk having an integral representation in it, *J. Soviet Math.*, 2(6)(1974), 606-617.

T. Hayami and S. Owa, The Fekete-Szegö problem for p-valently Janowski starlike and convex functions, *Int. J. Math. Math. Sci.*, (2011), Article Id. 583972.

W. Janowski, Some extremal problems for certain families of analytic functions, *Ann. Pol. Math.*, 28(1973), 297-326.

W. Kaplan, Close-to-convex schlicht functions, *Mich. Math. J.*, 1(1952), 169-185.

S. R. Keogh and E. P. Merkes, A coefficient inequality for certain classes of analytic functions, *Proc. Amer. Math. Soc.*, 20(1969), 8-12.

J. Krzyz, On the derivative of close-to-convex functions, *Colloq Math.*, 10(1963), 139-142.

R. J. Libera, Some radius of convexity problems, *Duke Math. J.*, 31(1964), 143-158.

M. S. Liu, On a subclass of p-valent close-to-convex functions of order β and type α, *J. Math. Study*, 30(1997), 102-104.

B. S. Mehrok, A subclass of close-to-convex functions, *Bull. Inst. Math. Acad. Sin.*, 10(4)(1982), 389-398.

B. S. Mehrok and G. Singh, A subclass of close-to-convex functions, *Int. J. Math. Anal.*, 4(27)(2010), 1319-1327.

Z. Nehari, Conformal Mappings, Mc. Graw Hill, New York, 1952.

K.I. Noor, On quasi-convex functions and related topics, *Int. J. Math. Math. Sci.*, 10(1987), 241-258.

S. Ogawa, A note on close-to-convex functions, *J. Nara Gakugei Univ.*, 8(1959), 9-10.

Y. Polatoglu, M. Bolkal, A. Sen and E. Yavuz, A study on the generalization of Janowski function in the unit disc, *Acta Mathematica Academiae Paedagogicae Nyiregyhaziensis*, 22(2006), 27-31.

M. O. Reade, On close-to-convex univalent functions, *Michigan Math. J.*, 3(1955-56), 59-62.

M. S. Robertson, On the theory of univalent functions, *Ann. Math.*, 37(1936), 169-185.

St. Ruscheweyh and T. Sheil-Small, Hadamard product of schlicht functions and the Polya-Schoenberg conjecture, *Comment. Math. Helv.* 48(1973), 119-135.

C. Selvaraj and S. Stelin, On a generalized class of quasi-convex functions, *Int. J. Pure Appl. Math.*, 101(6)(2015), 1051-1061.

C. Selvaraj, S. Stelin and S. Logu, On a subclass of close-to-convex functions associated with fixed second coefficient, *Appl. Comput. Math.*, 4(5)(2015), 342-345.

H. Singh and B. S. Mehrok, Subclasses of close-to-convex functions, *Tamkang J. Math.*, 44(4)(2013), 377-386.

G. Singh and G. Singh, Generalized subclasses of quasi-convex functions defined with subordination, *Elec. J. Math. Anal. Appl.*, 9(2)(2021), 172-185.

T. Umezawa, Multivalently close-to-convex functions, *Proc. Amer. Math. Soc.*, 8(1957), 869-874.

L. Xiong and X. Liu, A general subclass of close-to-convex functions, *Kragujevac J. Math.*, 36(2)(2012), 251-260.

Gagandeep Singh
Department of Mathematics
Khalsa College, Amritsar-143002, Punjab, India
Email: kamboj.gagandeep@yahoo.in

Gurcharanjit Singh
Department of Mathematics
G.N.D.U. College, Chungh(Punjab), India.

CORRECTION*
A remark on Normalized Laplacian eigenvalues of signed graph

B. PRASHANTH, K. NAGENDRA NAIK AND R. SALESTINA M

Abstract

With this article in mind, we have found some results using eigenvalues of graph with sign. It is intriguing to note that these results help us to find the determinant of *Normalized Laplacian* matrix of signed graph and their coefficients of characteristic polynomial using the number of vertices. Also we found bounds for the lowest value of eigenvalue.

Mathematics Subject Classification 2010: 05C50, 15A18, 05C22
Keywords: Marked graph; Signed graph; Switched signed graph Balanced signed graph.

1. INTRODUCTION

The readers should refer to [6] for expression and notations of graph theory and only simple and finite graphs are considered.
A Signed graph $\Gamma = (G(V, E), \nabla)$ is a graph with positive and negative signs in every edge, where G is the underlined graph without signs and ∇ is the function from the collection of edges E to the set having positive and negative signs.
One of the main applications of signed graphs is to represent the relationship among people where we assign a positive sign if the relationship between two individuals is pleasant, otherwise we assign a negative sign. [10] & [5].

The *balanced* signed graph was introduced by F.Harary[7] and he defines that every cycle of a *balanced* signed graph has negative edges in even number if not Γ is said to be *unbalanced*. In [8], Harary and Kebel showed a simple algorithm for balancing of a signed graph.

A graph that has been marked Γ_ν is a signed graph with positive or

* Correction of the title of the paper B. Prashanth, K. Nagendra Naik and R. Salestina M, Restructured class of estimators for population mean using an auxiliary variable under simple random sampling scheme. Journal of Applied Mathematics, Statistics and Informatics 17 (2), 75-90, 2021. DOI: https://doi.org/10.2478/jamsi-2021-0010

10.2478/jamsi-2022-0008

negative signs assigned to its vertices. The process of assigning signs to the vertices is called marking ν. For $v \in V(\Gamma)$, marked graph Γ_ν is defined as

$$\nu(v) = \prod_{uv \in E(\Gamma)} \nabla(uv).$$

Switched signed graph $\Gamma_\nu(\Gamma)$ was defined by R.P Abelson and Rosenburg [14] which paved the way for the study of social behavior and mathematical analysis in graph theory.

A signed graph Γ_2 is obtain from a signed graph Γ_1 by reversing the sign of edges of Γ_1 whose end vertices are having opposite sign, and their underlined graphs G_1 and G_2 are isomorphic. The signed graph Γ_1 *switching equivalent* to Γ_2, is represented as $\Gamma_1 \sim \Gamma_2$.

Following is the characterization of *switched signed graphs.*

PROPOSITION 1. [15] Any two signed graphs whose underlying graphs are same are cycle isomorphic if, and only if they are switching equivalent.

In a signed graph, degree of each vertex can be calculated by $d = d^+ + d^-$ so that degree of vertices in a signed graph Γ and their underlined graph is the same.

In *adjacent* matrix $A(\Gamma)$, if two vertices are adjacent then the entry a_{ij} is 1 along with the sign of the edge, otherwise the entry is zero.

In a *Laplacian* matrix $L(\Gamma)$, if vertices v_i and v_j are adjacent then the entry a_{ij} is 1 with the opposite sign of corresponding adjacent edge v_iv_j, otherwise a_{ij} is zero and the diagonal entries a_{ii} being the degree of the vertex . Also $L(\Gamma)$=$S(\Gamma) - A(\Gamma)$, where S is the diagonal matrix.

Here $(\Gamma, -)$ is a signed graph in which each edge is assigned by minus sign and $L(\Gamma, -)$ is the *Laplacian* matrix of $(\Gamma, -)$. Eigenvalues of *Laplacian* matrix of a signed graph are $\lambda_1 \geq \lambda_2 \geq \lambda_3 \; .... \geq \lambda_n$.

2. *NORMALIZED LAPLACIAN* MATRIX OF SIGNED GRAPH

F.R.K. Chung [17] introduced the *Normalized Laplacian* matrix. Lower bounds of *Normalized Laplacian* were investigated by Grossman [13] and its properties were stated by Chen et. al. in [11]. Also in [4] Cvetkovic et.al. mentioned deeply about normalized Laplacian and their bounds of eigen values.

The *Normalized Laplacian* matrix $\mathsf{L}(\Gamma)$ of a signed graph Γ with vertices u and v is given by

$$\mathsf{L}_{uv} = \begin{cases} 1, \text{ if } u = v \textit{ and } d_u \neq 0 \\ -\nabla(uv)\frac{1}{\sqrt{d_u d_v}}, \text{ if } u \text{ and } v \text{ are adjacent} \\ 0 \text{ otherwise.} \end{cases}$$

Let $0 \leq \mu_1 \leq \mu_2 \leq \mu_3 \; .... \leq \mu_n$ be the eigenvalues of *Normalized Laplacian* matrix of Γ, with n vertices. Also $\mathsf{L}(\Gamma) = S^{-1/2}L(\Gamma)S^{-1/2}$.

In 2003 Yaoping Hou. et. al. [16] established new bounds in the following theorem.

THEOREM 2. [16] Let Γ be a signed graph with n vertices. Then

$$\lambda_1 \leq 2(n-1),$$

equality applies if and only if Γ is switching equivalent to a complete graph with all edges being negative.

Some of the novel results prompted by the above theorem are presented in this article.

THEOREM 3. Let $\Gamma = (G, \nabla)$ be a signed graph. The greatest eigenvalue of *Normalized Laplacian* matrix $\mathsf{L}(\Gamma)$ is 2 if and only if Γ is switching equivalent to a complete graph with all edges being negative.

PROOF. If $\Gamma \sim (K_n, -)$ then $\mu_n(\mathsf{L}(\Gamma)) = \mu_n(S^{-1/2}\; L(\Gamma)\; S^{-1/2})$

$$=\mu_n(S^{-1/2}\; (S(\Gamma) - A(\Gamma))\; S^{-1/2})$$

$$=\mu_n(S^{-1/2}\; (S(\Gamma))S^{-1/2} - S^{-1/2}(A(\Gamma))\; S^{-1/2})$$

$$=\mu_n(S^{-1/2}\ (S(\Gamma))S^{-1/2}) - \mu_n(S^{-1/2}(A(\Gamma))\ S^{-1/2})$$

$$=\mu_n(S^{-1/2}\ (S(\Gamma))S^{-1/2}) + \mu_n(S^{-1/2}(-A(\Gamma))\ S^{-1/2})$$

$$= 1 + 1$$

$$= 2.$$

If $\mu_n = 2$ *then* $\mu_n(S^{-1/2}(S(\Gamma)S^{-1/2}) = \mu_n(S^{-1/2}(-A(\Gamma))S^{-1/2})$.

Thus, $\mu_n(A(\Gamma)) = \mu_n(-(J-I))$, *where* J *is the all one matrix.*

Hence $\Gamma \sim (K_n, -)$.

THEOREM 4. Let Γ be a signed graph and K_n be the complete graph with n vertices, $\Gamma \sim (K_n, -)$ if and only if $\mu_k = \frac{n-2}{n-1}$, for $k < n$.

PROOF. *If* $\Gamma \sim (K_n, -)$ then $\mu_k(\mathsf{L}(\Gamma)) = \mu_k(S^{-1/2}(L(\Gamma))\ S^{-1/2})$

$$=\mu_k(S^{-1/2}\ (S(\Gamma) - A(\Gamma))\ S^{-1/2})$$

$$=\mu_k(S^{-1/2}\ (S(\Gamma))S^{-1/2} - S^{-1/2}(A(\Gamma))\ S^{-1/2})$$

$$=\mu_k(S^{-1/2}\ (S(\Gamma))S^{-1/2}) - \mu_k(S^{-1/2}(A(\Gamma))\ S^{-1/2})$$

$$=\mu_k(S^{-1/2}\ (S(\Gamma))S^{-1/2}) + \mu_k(S^{-1/2}(-A(\Gamma))\ S^{-1/2})$$

$$= 1 + \frac{-1}{n-1}$$

$$= \frac{n-2}{n-1}.$$

If $\mu_k = \frac{n-2}{n-1}$ *then* $\mu_k(S^{-1/2}\ (S(\Gamma))S^{-1/2}) = \mu_k(S^{-1/2}(-A(\Gamma))S^{-1/2})$.

Thus, $\mu_k(S^{-1/2}\ A(\Gamma)\ S^{-1/2}) = \mu_k\ (S^{-1/2}(-(J-I))\ S^{-1/2})$, *where* J *is the* *all one matrix. Hence* $\Gamma \sim (K_n, -)$.

COROLLARY 5. Let Γ be a signed graph. The greatest eigenvalue of *Normalized Laplacian* matrix $\mathsf{L}(\Gamma)$ is 2 if and only if Γ is switching equivalent to a complete bipartite graph with all edges being negative.

PROPOSITION 6. Let Γ be a graph with sign. If $\Gamma \sim (K_n, -)$ then $\sum_{i=1}^{n} \mu_i = n$.

PROOF. $\sum_{i=1}^{n} \mu_i = \mu_1 + \mu_2 + \mu_3 + ... + \mu_n$

$= 2 + (n-1)\ \frac{n-2}{n-1}$

$= 2 +$ n - 2

$=$ n.

3. DETERMINANT OF *NORMALIZED LAPLACIAN* MATRIX OF SIGNED GRAPH

3.1. Matrix Tree Theorem for a *Laplacian* matrix

If b_c be the number of essential spanning subgraphs which contain c negative cycles, then

$$Det(L(\Gamma)) = \sum_{c=0}^{n} 4^c\ b_c.$$

From the above matrix tree theorem, we determine the determinant of *Normalized Laplacian* matrix of a graph with sign and n number of vertices.

PROPOSITION 7. Let Γ be a signed graph. If $\Gamma \sim (K_n, -)$ then

$$Det(L(\Gamma)) = 2\{\tfrac{(n-2)}{(n-1)}\}^{(n-1)}\ .$$

PROOF.

$$Det(\mathsf{L}(\Gamma)) = \prod_{i=1}^{n} \mu_i$$

$$= 2 \cdot \{1 - \tfrac{1}{n-1}\} \cdot \{1 - \tfrac{1}{n-1}\} \cdots \{1 - \tfrac{1}{n-1}\}$$

$$= 2 \cdot \tfrac{n-2}{n-1} \cdot \tfrac{n-2}{n-1} \cdot \tfrac{n-2}{n-1} \cdots \tfrac{n-2}{n-1}$$

$$= 2 \cdot \{\tfrac{n-2}{n-1}\}^{(n-1)} .$$

4. CHARACTERISTIC POLYNOMIAL COEFFICIENTS OF A *NORMALIZED LAPLACIAN* MATRIX

In the study of chemical properties of molecules and their bond structures, coefficients of a characteristic polynomial play a vital role. Ivailo M. Mladenov et. al. [12] introduced an algorithm to find the coefficients of characteristic polynomial of adjacent matrix of a graph. Kel'man expanded the latter formula and it is known as Kel'man formula.

Kel'man formula is the method to find the coefficients of a characteristic polynomial of a matrix which is given as follows.

THEOREM 8. [1]

Let G be a simple graph. Then the characteristic polynomial coefficients of a *Normalized Laplacian* matrix of the graph are provided by using

$$b_{n-k} = (-1)^{n-k} \textstyle\sum_{F \in F_k} \gamma(F) \qquad \text{where } k \geq 1 \text{ (for } k = 0, b_n = 0.)$$

F_k denotes the set of forests in G having k components and

$$\gamma(G) = \prod_{i=1}^{k} |F_i|$$

is the product of the orders of the components of the forest F.

Also in [3] Carla Silva Oliveria et. al. have found second and third *Laplacian* coefficients of a characteristic polynomial in 2002. Francesco Belardo and Slobodan K. Simic [9] have found *Laplacian* coefficients of signed graph by the following theorem:

THEOREM 9. [9] The *Laplacian* characteristic polynomial of Γ is given by $\psi(\Gamma, x) = x^n + b_1 x^{n-1} + ... + b_{n-1}x + b_n$ for any signed graph Γ, then

$$b_i = (-1)^i \sum_{H \in H_i} w(H)$$

where H_i denotes the set of signed TU- subgraphs of Γ containing i edges.

We now present a simplified way of finding the coefficient of *Normalized Laplacian matrix* using the number of vertices.

PROPOSITION 10. Let $\Gamma = (G, \nabla)$ be a signed graph. If $\Gamma \sim (K_n, -)$, then for a positive integer t,

$$tr(\mathsf{L}^t) = 2^t + \{\frac{(n-2)^t}{(n-1)^{(t-1)}}\}$$

.

PROOF.
$tr(\mathsf{L}) = \sum_{i=1}^{n} \mu_i$

$$= 2 + (n-1)\tfrac{(n-2)}{(n-1)}$$

$$tr(\mathsf{L}^2) = 2^2 + (n-1)\{\tfrac{n-2}{n-1}\}^2.$$

$$tr(\mathsf{L}^3) = 2^3 + (n-1)\{\tfrac{n-2}{n-1}\}^3.$$

Similarly for an integer k,

$$tr(\mathsf{L}^k) = 2^k + (n-1)\{\tfrac{n-2}{n-1}\}^k$$

$$tr(\mathsf{L}^{k+1}) = \textstyle\sum_{i=1}^{n} \mu_i^{k+1}$$

$$= 2^{k+1} + \{\tfrac{n-2}{n-1}\}^{k+1} + \ldots + \{\tfrac{n-2}{n-1}\}^{k+1}$$

$$= 2^{k+1} + (n-1)\{\tfrac{n-2}{n-1}\}^{k+1}.$$

Hence by induction,

$$tr(\mathsf{L}^t) = 2^t + (n-1)\{\tfrac{n-2}{n-1}\}^t$$

$$tr(\mathsf{L}^t) = 2^t + \{\tfrac{(n-2)^t}{(n-1)^{(t-1)}}\}.$$

Examples:

$tr(\mathsf{L}) = 2 + (n-1)\{\frac{n-2}{n-1}\} = n.$

$tr(\mathsf{L}^2) = 2^2 + (n-1)\{\frac{n-2}{n-1}\}^2 = \frac{n^2}{n-1}.$

$tr(\mathsf{L}^3) = 2^3 + (n-1)\{\frac{(n-2)^3}{(n-1)^3}\} = \frac{n^3+2n^2-4n}{(n-1)^2}.$

$tr(\mathsf{L}^4) = 2^4 + (n-1)\{\frac{(n-2)^4}{(n-1)^4}\} = \frac{n^4+8n^3-24n^2+16n}{(n-1)^3}.$

Coefficients of characteristic polynomial of a *Normalized Laplacian* matrix of signed graph Γ, a_1, a_2, a_3, a_4 are calculated as follows.

$$a_1 = -tr(\mathsf{L}) = -n.$$

$$a_2 = \frac{-1}{2} tr(B_1\mathsf{L}) \quad \text{where } B_1 = \mathsf{L} + a_1 I$$

$$= \frac{-1}{2}(tr(\mathsf{L}^2) - ntr(\mathsf{L}))$$

$$= \frac{1}{2}\{\frac{n^2(n-2)}{(n-1)}\}.$$

$$a_3 = \frac{-1}{3} tr(B_2\mathsf{L}) \quad \text{where } B_2 = B_1\mathsf{L} + a_2 I$$

$$= \frac{-1}{3} tr(B_1\mathsf{L}^2 + a_2\mathsf{L})$$

$$= \frac{-1}{3}(tr(\mathsf{L}^3) + a_1 tr(\mathsf{L}^2) + a_2 tr(\mathsf{L}))$$

$$= \frac{-1}{3}\{8 + \frac{(n-2)^3}{(n-1)^2} - n(4 + \frac{(n-2)^2}{(n-1)}) + \frac{n}{2}(\frac{n^2(n-2)}{(n-1)})\}$$

$$= \frac{-1}{6}(\frac{n^5-5n^4+6n^3+4n^2-8n}{(n-1)^2}).$$

$$a_4 = \frac{-1}{4} tr(B_3\mathsf{L}) \quad \text{where } B_3 = B_2\mathsf{L} + a_3 I$$

$$= \frac{-1}{4} tr(B_2\mathsf{L}^2 + a_3\mathsf{L})$$

$$= \frac{-1}{4} tr((B_1 \mathsf{L} + a_2)\mathsf{L}^2 + a_3 \mathsf{L})$$

$$= \frac{-1}{4} tr((\mathsf{L} + a_1)\mathsf{L}^3 + a_2 \mathsf{L}^2 + a_3 \mathsf{L})$$

$$= \frac{-1}{4} (tr(\mathsf{L}^4) + a_1 tr(\mathsf{L}^3) + a_2 tr(\mathsf{L}^2) + a_3 tr(\mathsf{L}))$$

$$= \frac{-1}{4} \left(\frac{n^4+8n^3-24n^2+16n}{(n-1)^3} - \frac{n^4+2n^3-4n^2}{(n-1)^2} + \frac{n^5-2n^4}{2(n-1)^2} - \frac{n^6-5n^5+6n^4+4n^3-8n^2}{6(n-1)^2} \right)$$

$$= \frac{1}{24} \left(\frac{n^7-9n^6+26n^5-8n^4-96n^3+176n^2-96n}{(n-1)^3} \right).$$

THEOREM 11. Let Γ be any signed graph which is switching equivalent to a complete graph in which each edge is negative and $\psi(\Gamma, x) = x^n + a_1 x^{n-1} + ... + a_{n-1} x + a_n$ be the *Normalized Laplacian* characteristic polynomial of Γ with $a_0 = 1$. Then,

$$a_\varsigma = \frac{-1}{\varsigma} \sum_{m=0}^{\varsigma-1} a_m tr(\mathsf{L}^{\varsigma-m})$$

where, a_ς is the coefficient of characteristic polynomial and $\varsigma \neq 0$.

PROOF.

Since $a_1 = -tr(\mathsf{L}) = -n$,

$a_2 = \frac{-1}{2} tr(B_1 \mathsf{L})$ *where* $B_1 = \mathsf{L} + a_1 I$

$= \frac{-1}{2} (a_0 tr(\mathsf{L}^2) + a_1 tr(L)).$

Similarly for an integer k,

$$a_k = \frac{-1}{k} (a_0 tr(\mathsf{L}^k) + a_1 tr(\mathsf{L}^{k-1}) + a_2 tr(\mathsf{L}^{k-2}) + ... + a_{k-1} tr(\mathsf{L}))$$

i.e., $a_k = \frac{-1}{k} \sum_{m=0}^{k-1} a_m tr(\mathsf{L}^{k-m})$.

$a_{k+1} = \frac{-1}{k+1} tr(B_k \mathsf{L})$ *where* $B_k = B_{k-1} \mathsf{L} + a_k I$

$$= \frac{-1}{k+1} tr(B_{k-1}\mathsf{L}^2 + a_k\mathsf{L})$$

$$= \frac{-1}{k+1} tr((\mathsf{L}B_{k-2} + a_{k-1})\mathsf{L}^2 + a_k\mathsf{L})$$

$$= \frac{-1}{k+1} tr(\mathsf{L}^3 B_{k-2} + a_{k-1}\mathsf{L}^2 + a_k\mathsf{L})$$

$$= \frac{-1}{k+1} (tr(\mathsf{L}^3 B_{k-2}) + a_{k-1}tr(\mathsf{L}^2) + a_k tr(\mathsf{L}))$$

$$a_{k+1} = \frac{-1}{k+1} (tr(\mathsf{L}^4 B_{k-3}) + tr(\mathsf{L}^3)a_{k-1} + tr(\mathsf{L}^2)a_k + tr(\mathsf{L}))$$

. . .

$$a_{k+1} = \frac{-1}{k+1} (a_0 tr(\mathsf{L}^{k+1}) + a_1 tr(\mathsf{L}^k) + a_2 tr(\mathsf{L}^{k-1}) + ... + a_k tr(\mathsf{L}))$$

i.e., $a_{k+1} = \frac{-1}{k+1} \sum_{m=0}^{k} a_m tr(\mathsf{L}^{k+1-m})$.

Hence by induction,

$$a_\varsigma = \frac{-1}{\varsigma} \sum_{m=0}^{\varsigma-1} a_m tr(\mathsf{L}^{\varsigma-m}) \qquad where\ \varsigma \neq 0.$$

Corollary 12. For any signed graph Γ and $\Gamma \sim (K_n, -)$, the *Normalized Laplacian* characteristic polynomial of Γ is $\psi(\Gamma, x) = a_0x^n + a_1x^{n-1} + ... + a_{n-1}x + a_n$ with $a_0 = 1$ then,

$$a_\varsigma = \frac{-1}{\varsigma} \sum_{m=0}^{\varsigma-1} a_m \{2^{\varsigma-m} + \frac{(n-2)^{\varsigma-m}}{(n-1)^{(\varsigma-(m+1))}}\} \qquad where, \varsigma \neq 0.$$

Proof. *By* Proposition 10,

$$tr(\mathsf{L}^t) = 2^t + \{\frac{(n-2)^t}{(n-1)^{(t-1)}}\}.$$

By Theorem 11,

$$a_\varsigma = \frac{-1}{\varsigma} \sum_{m=0}^{\varsigma-1} a_m tr(\mathsf{L}^{\varsigma-m}) \qquad where, \varsigma \neq 0.$$

i.e.,

$$a_\varsigma = \frac{-1}{\varsigma} \sum_{m=0}^{\varsigma-1} a_m \{2^{\varsigma-m} + \frac{(n-2)^{\varsigma-m}}{(n-1)^{(\varsigma-(m+1))}}\}.$$

COROLLARY 13. any signed graph Γ and $\Gamma \sim (K_n, -)$, the *Normalized Laplacian* characteristic polynomial of Γ is $\psi(\Gamma, x) = a_0 x^n + a_1 x^{n-1} + ... + a_{n-1} x + a_n$ with $a_0 = 1$ and μ_k be the *Normalized Laplacian* eigenvalue of Γ then,

$$a_\varsigma = \frac{-1}{\varsigma} \sum_{m=0}^{\varsigma-1} a_m (2^{\varsigma-m} + (n-1)\mu_k^{\varsigma-m})$$

where, $\varsigma \neq 0$ and $k < n$.

PROOF. *From* theorem 4, $\mu_k = \frac{(n-2)}{(n-1)}$ hence by corollary 12,

$$a_\varsigma = \frac{-1}{\varsigma} \sum_{m=0}^{\varsigma-1} a_m (2^{\varsigma-m} + (n-1)\mu_k^{\varsigma-m})$$

where, $\varsigma \neq 0$ and $k < n$.

5. BOUNDS OF EIGENVALUES OF *NORMALIZED LAPLACIAN* MATRIX OF SIGNED GRAPH.

As we know, $\mathsf{L} = S^{-1/2} L S^{-1/2}$, where $S^{-1/2}$ is invertible.
The vectors g and g_j are defined as:
If f is the eigen function of L corresponding to eigenvalue μ_k, then $g = U^{1/2} f$,

$g_j = U^{1/2}\ f_j$

$$\mu_k = inf \frac{\sum (f(u) - \nabla(u,v) f(v))^2}{\sum_u f^2(u) d_u}$$

where degree of the vertex u is d_u.

If $\Gamma \sim (K_n, -)$ and for a vertex v

$$(1-\mu_k)f(v) = \frac{1}{d_v}\sum_{u\sim v}\nabla(u,v)f(u)$$

where, $u \sim v$ means u and v vertices are adjacent. Let $v_1, v_2, ..., v_{m+1}$ be adjacent vertices sequence, $f(v_1)$ be maximal and $f(v_{m+1}) \leq 0$. Let $y_i = f(v_i)$ and $\beta = 1-\mu_k$. We get

$$\beta y_1 = \frac{1}{d_{v_1}}\sum_{u\sim v_1}\nabla(u,v_1)f(u) \leq \frac{y_2}{d_{v_1}} + \frac{(deg(v_1)-1)y_1}{d_{v_1}} \leq \frac{y_2}{d} + \frac{(d-1)y_1}{d}$$

.

Assume $y_1 = 1$, so that $y_2 \geq 1-\mu_k d$.

In view of the fact, v_i is adjacent to v_{i+1} and v_{i-1} for $2 \leq i \leq m$, we get

$$\beta y_i \leq \frac{y_{i-1}+y_{i+1}}{d} + \frac{(d-2)y_1}{d}.$$

This implies

$$y_{i+1} \geq \beta d y_i - y_{i-1} - (d-2).$$

If $\Gamma \sim (K_n, -)$ we observe the following:

(1) $(1-\mu_k) = \frac{1}{(n-1)}$

(2) $f(v_1) = \sum_{u\sim v_1}\nabla(u,v_1)f(u)$

(3) $y_2 \geq 3-n$

(4) $y_{m+1} \geq 3-2n$.

Proposition 14. For $3 \leq r \leq m+1$, $y_r \geq 1-\mu_k\beta^{r-3}d^{r-2}-\mu_k\beta^{r-2}d^{r-1}$.

Proof. *We have,*

$$y_2 \geq 1-\mu_k d \tag{1}$$

$$y_{i+1} \geq \beta d y_i - y_{i-1} - (d-2). \tag{2}$$

Proof is by induction on r.

From (2),

$$y_3 \geq 1 - \mu_k d - \mu k \beta d^2.$$

Suppose result holds for $r \leq i$, *where* $i \geq 3$.

From (2)

$$y_3 \geq \beta d y_2 - 1 - (d-2)$$

$$y_4 \geq \beta d y_3 - y_2 - (d-2)$$

$$y_5 \geq \beta d y_4 - y_3 - (d-2)$$

$$\cdot \quad \cdot \quad \cdot$$

$$y_i \geq \beta d y_{i-1} - y_{i-2} - (d-2)$$

$$y_{i+1} \geq \beta d y_i - y_{i-1} - (d-2).$$

$\therefore \ (y_2 + y_3 + y_4 + ... + y_i + y_{i+1}) \geq \beta d(y_2 + y_3 + y_4 + ... + y_{i-1} + y_i) - (y_1 + y_2 + y_3 + y_4 + ... + y_{i-2} + y_{i-1}) - (i-1)(d-2) + 1 - \mu_k d.$

$y_{i+1} \geq \beta d(y_2+y_3+y_4+...+y_{i-1}+y_i) - (2y_2+2y_3+2y_4+...+2y_{i-1}+2y_i) - (i-1)(d-2) + y_i - \mu_k d.$

i.e., $$y_{i+1} \geq (\beta d-2)(y_2+y_3+y_4+...+y_{i-1}+y_i) + y_i - (i-1)(d-2) - \mu_k d \quad (3)$$

From (2) we have,

$$y_3 \geq 1 - \mu_k d - \mu_k \beta d^2$$

$$y_4 \geq 1 - \mu_k \beta d^2 - \mu_k \beta^2 d^3.$$

In general

$$y_i \geq 1 - \mu_k \beta^{i-3} d^{i-2} - \mu_k \beta^{i-2} d^{i-1}$$

so,

$(y_2+y_3+y_4+...+y_{i-1}+y_i) \geq (i-1)-2\mu_k d-2\mu_k\beta d^2-2\mu_k\beta^2 d^3...2\mu_k\beta^{i-3}d^{i-2}-\mu_k\beta^{i-2}d^{i-1}$.

Also we have

$$y_i \geq 1-\mu_k\beta^{i-3}d^{i-2}-\mu_k\beta^{i-2}d^{i-1}.$$

From (3) we get $y_{i+1} \geq (\beta d-2)((i-1)-2\mu_k d-2\mu_k\beta d^2-2\mu_k\beta^2 d^3-...-2\mu_k\beta^{i-3}d^{i-2}-\mu_k\beta^{i-2}d^{i-1})+(1-\mu_k\beta^{i-3}d^{i-2}-\mu_k\beta^{i-2}d^{i-1})-\mu_k d-(i-1)(d-2)$

$y_{i+1} \geq 1-(i-4)\mu_k d+2\mu_k\beta d^2+2\mu_k^2 d^3...2\mu_k\beta^{i-4}d^{i-3}+\mu_k\beta^{i-3}d^{i-2}-\mu_k\beta^{i-1}d^i-\mu_k\beta^{i-2}d^{i-1}$

.

$$y_{i+1} \geq 1+(i-3)\mu_k d+2\mu_k d+2\mu_k d+...+2\mu_k d+\mu_k d-\mu_k\beta^{i-2}d^{i-1}-\mu_k\beta^{i-1}d^i$$

$$y_{i+1} \geq 1+(i-3)\mu_k d-\mu_k\beta^{i-2}d^i-\mu_k\beta^{i-2}d^{i-1}$$

$$y_{i+1} \geq 1-\mu_k\beta^{i-1}d^i-\mu_k\beta^{i-2}d^{i-1}.$$

As a result, $y_r \geq 1-\mu_k\beta^{r-3}d^{r-2}-\mu_k\beta^{r-2}d^{r-1}$.

THEOREM 15. Let Γ be a graph with sign having n vertices and if $\Gamma \sim (K_n,-)$. Then $\mu_k \geq \frac{1}{n}$.

PROOF. *From* Proposition 14,

$$0 \geq y_{m+1} \geq 1-\mu_k\beta^{m-2}d^{m-1}-\mu_k\beta^{m-1}d^m$$

$$0 \geq 1-\mu_k d^{m-1}-\mu_k d^m$$

$$\mu_k \geq \frac{1}{(d+1)d^{m-1}}.$$

The distance between a vertex that maximises f and one that minimises f is at most the graph's diameter 'D'., therefore $m \leq \lceil D/2 \rceil$. *Since the diameter is 1,*

$$\mu_k \geq \frac{1}{(d+1)}.$$

Hence the result follows.

6. CONCLUSION

Usually the coefficients of the characteristic polynomial of a graph or a signed graph are found using the concept of trees and TU subgraphs. But in this paper, we have given a simple and an elegant proof of finding the *Laplacian* coefficients of the characteristic polynomial of a signed graph using the number of vertices of the graph. We believe that this new approach will pave way for further research in this area.

REFERENCES

A. K. Kel'mans, Properties of the characteristic polynomial of a graph, *Kibernetiky na Slŭ Kommunizmu*, Emergija, Markova-Leningrand, 4(1967), 27-41(in Russian).

B. Prashanth, K. Nagendra Naik, K. R. Rajanna, A remark on eigen values of Signed graph, *Journal of Applied Mathematics, Statistics and Informatics*15(1) (2019), 33-42. (ISSN (print): 1336-9180,(ISSN (on-line): 1339-0015, Poland)).

C. S. Oliveria, N. M. Maia de Abreu, S. Jurkiewicz, The Characteristic Polynomial of the Laplacian of graphs in(a,b)-linear classes, *Linear Algebra and its Applications*. 356,113-121 (2002), (ISSN:0024-3795).

D. Cvetkovic, P. Rowlinson, S. Simi'c: An Introduction to the Theory of Graph Spectra. Cambridge University Press, Cambridge(2010).

F. Harary, R. Z. Norman and D. W. Cartwright, *Structural models: An introduction to the theory of directed graphs*, Wiley Inter-Science, Inc., New York, 1965.

F. Harary, Graph Theory, Addison-Wesley Publishing Co., 1969.

F. Harary, On the notion of balance of a signed graph, *Michigan Math. J.*, 2(1953), 143-146.

F. Harary and J. A. Kabell, Counting balanced signed graphs using marked graphs, *Proc. Edinburgh Math. Soc.*, 24 (2)(1981), 99-104.

F. Belardo and S. K. Simic, On the Laplacian coefficients of signed graphs, *Linear ALgebra and its Applications,*475 (2015), 94-113. (ISSN:0024-3795).

G. T. Chartrand, *Graphs as Mathematical Models*, Prindle, Weber & Schmidt, Inc., Boston, Massachusetts 1977.

G. Davis, G. Chen, F. Hall, Z. Li, M. Stewart and K. Patel, An interlacing result on normalized Laplacian, *SIAM J. Discrete Math.* 18(2004),353-361.

I. M. Mladenov, M. D. Kotarov and J. G. Vassileva Popova, Method for Computing the Characteristic Polynomial, *International Journal of Quantum Chemistry.* 10(8),339-341 (1980), (ISSN:0020-7608-80).

J. P. Grossman, An eigenvalue bound for the Laplacian of a graph, *Discrete Math.* 300 (2005), 225-228.

M. J. Rosenberg and R. P. Abelson, Symbolic psycho-logic: A model of attitudinal cognition, *Behav. Sci.*, volume 3 , p. 1-13(1958).

T. Zaslavsky, Signed graphs, *Discrete Appl. Math.*, 4(1)(1982), 47-74.

Y. Hou , Y. Pan and J. Li On the Laplacian Eigenvalues of signed graphs, *Linear and Multilinear Algebra*,(2003) p. 21-30.

F. R. K. Chung, Counting balanced signed graphs using marked graphs, *Proc. Edinburgh Math. Soc.*, 24 (2)(1981), 99-104.

B. Prashanth
Department of Mathematics,
JSS Science and Technology University
Mysuru-570 006, India.

K. Nagendra Naik
Department of Mathematics,
Atria Institute of Technology
Hebbal
Bangalore-560 024, India.

Ruby Salestina M
Department of Mathematics,
Yuvaraja's College
University of Mysore
Mysuru, India.

Corresponding author: bprassy@gmail.com