Received 8 November 2024, accepted 10 December 2024, date of publication 13 December 2024, date of current version 23 December 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3517321

RESEARCH ARTICLE

A Stochastic Residual-Based Dominant Component Analysis for Speech Enhancement

RABIA SHARIF[1], SHAZIA JAVED[1], MUSAED ALHUSSEIN[2], UZMA BASHIR[1], KHURSHEED AURANGZEB[2], (Senior Member, IEEE), AND BHARAT BHUSHAN[3]

[1]Department of Mathematics, Lahore College for Women University, Lahore 54600, Pakistan
[2]Department of Computer Engineering, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia
[3]Department of Computer Science and Engineering, School of Engineering and Technology, Sharda University, Greater Noida 201310, India

Corresponding author: Shazia Javed (Shazia.Javed@lcwu.edu.pk)

This work was supported by King Saud University, Riyadh, Saudi Arabia, through the Researchers Supporting Project under Grant RSPD2024R553.

ABSTRACT Noise and sparsity often affect speech signals, leading to serious problems in processing and communication. Speech enhancement is required to improve the quality of the speech signals. This paper introduces a new technique that combines a stochastic approach and dominant component analysis, a variant of principal component analysis for adaptive data analysis. The stochastic approach is a modeling technique that takes into account uncertainty and random fluctuations in the signal. This allows for a more precise estimation of residuals. The proposed method involves estimating residuals using a stochastic approach, which subsequently accumulate into a matrix. Adaptively, we compute the dominant components of the residual matrix. We then use these components to reconstruct clean, enhanced speech. The proposed method aims to forecast sparse data, eliminate noise, and minimally affects crucial data attributes such as energy, covariance, dynamic range, and RMS amplitude.

INDEX TERMS Adaptive PCA, affine projection, noise cancellation, sparsity removal, speech enhancement.

I. INTRODUCTION

Speech enhancement can be defined as improving the quality and clarity of the speech signals degraded by noise, reverberation, or any other factors. It is essential in a variety of applications, including telecommunications, hearing aids, voice commands, control, and automatic speech recognition. The speech enhancement techniques aim to isolate the target speech from the background noise or interference while preserving the desirable qualities of the speech [1].

There are many ways to improve speech, such as spectral subtraction, Wiener filtering [2], statistical model-based methods, deep neural networks, and beamforming. These techniques aim to enhance speech clarity and intelligibility by reducing noise and interference. Several deep learning models for speech enhancement employ mapping-based techniques [3], while others use mask-based methods [4], [5]. In the past, there have been extensive efforts to conduct a survey in the field of speech. The survey article authored by Zhang et al. [6] offers a comprehensive review of pertinent deep learning methods, primarily focusing on their application to noise-resistant speech recognition tasks. Michelsanti et al. [7] conducted a study on deep learning techniques for improving and separating audio-visual speech. The paper by Das et al. [8] provides an introduction of the basic principles of speech enhancement. However, it specifically focuses on discussing strategies based on deep learning. Lin et al. [9] employ a multi-stage learning strategy to increase speech by applying a sequence of TCM blocks. The network specifically builds each block with a self-attentive TCM to progressively improve the magnitude spectrogram. Pandey and Wang [10] introduce a dense

The associate editor coordinating the review of this manuscript and approving it for publication was Fang Yang.

convolutional network (DCN) structure enhanced with self-attention. Skip connections enhance the encoder-decoder arrangement in the architecture. Both the encoder and decoder contain layers that consist of a dense block combined with an attention module. This design utilizes dense blocks and attention modules to efficiently extract features by prioritizing feature reuse, increasing network depth, and optimizing context aggregation. Tan and Wang [11] propose two compression techniques to reduce the size of deep neural network (DNN) models for speech enhancement. These techniques involve sparse regularization, iterative pruning, and quantization using clustering. Zhang et al. [12] proposed a weighted magnitude-phase loss function to emphasize the need for precise magnitude estimates. Poluboina et al. [13] have devised a method to diminish noise by specifically targeting the reduction of the mean square error (MSE) between the temporal envelopes of the improved speech and the clean speech. This methodology is particularly well-suited for cochlear implant (CI) applications.

Kumar et al. [14] utilize a semi-soft thresholding method to eliminate noise and obtain a clear signal. Surbhi Bharti conducted an evaluation of the techniques designed to decrease noise and improve voice quality in order to meet the needs of forensic audio analysis, which can be used as evidence [15]. Abajaddi and her colleagues proposed the use of a hybrid approach involving single-frequency filtering (SFF) and a modified spectral subtraction technique to enhance single-channel speech [16]. Bipasha and her colleagues [17] employ a combination of Siamese architecture and hybrid RNN models to address the issue of noise reduction. To find out how well different adaptive algorithms worked, Kasukurthy [18] looked at their signal-to-noise ratio (SNR), mean square error (MSE), and root mean square error (RMSE) values. Chhetri [19] offers an in-depth analysis of speech enhancement strategies, discussing the difficulties and uses in different domains. This review aims to assist in choosing appropriate approaches to enhance speech quality and clarity. Brandon Colelough [20] investigates how the sampling rate of training data affects the performance of lightweight and efficient deep neural networks (DNNs) designed to work within the limitations of mobile devices. Among these, residual-based techniques have gained significant attention due to their effectiveness in preserving the natural quality of speech while reducing noise. The residual-based method focuses on the residual components of the speech signal, which contain essential features often masked by noise. These techniques can significantly improve speech quality and intelligibility by accurately estimating and enhancing residuals, making them highly valuable in environments with challenging acoustic conditions Signal and image/video processing frequently employ residual-based approaches for noise reduction or enhancement applications. The goal is to divide the input signal into two components: expected or estimated and residual. The anticipated component depicts the signal's underlying structure or model, whereas the residual component captures the signal's deviations. It is possible to improve the signal's quality or extract important information by analyzing and modifying the residuals. Applications such as image de noising, picture super-resolution, video compression, object recognition, and speech enhancement have proven the usefulness of residual-based approaches. They provide a framework for modeling and manipulating residuals, resulting in improved performance and better data utilization.

The stochastic gradient descent (SGD) technique is one of the premier methods utilized in machine learning for optimizing cost functions and computing residuals, essentially improving the efficiency and accuracy of learning models. Algorithms like the least mean square (LMS) and the affine projection algorithm (APA) are also recognized for their stochastic properties. Adaptive filters particularly value the LMS algorithm for its simplicity and effectiveness. It adjusts the filter coefficients iteratively to minimize the error between the desired and the actual output. However, the APA, which is an extension of the LMS [21], is praised for achieving faster convergence in settings with large eigenvalue spreads. This makes it ideal for uses that need to adapt quickly.

Adaptive dominant component analysis is a variation of the classic PCA technique intended to manage constantly changing data environments. Adaptive PCA continuously updates the main components as new data arrives, unlike static PCA, which typically operates on a fixed data set. It is appropriate for applications that collect data over time, such as online learning, real-time monitoring, and dynamic system analysis. Online Principal Component Analysis has revolutionised data analysis in environmental monitoring [22], finance [23], and network security [24]. Adaptive principal components can be used to lessen the effect of outliers and do feature selection in fields like image processing, text classification, and image processing [25]. To find multi-degree tooth-cut failures in gearboxes that are running at different speeds, fault diagnostic models can use adaptive noise canceling technique (ANCT) and distance ratio principal component analysis (DRPCA) [26]. Adaptive principal components can be used in fields like college admission, healthcare, and credit approval to ensure fairness and robustness of the learned projection [27].

DCA has been addressed using a variety of algorithms and methodologies, such as stochastic gradient descent and recursive formulations [28]. LMS and modified LMS algorithms frequently get preference over RLS in various applications due to their resilience in dealing with non-stationary data, computational speed, ability to withstand noise and outliers, efficient memory usage, suitability for online and real-time processing, adaptability to different signal characteristics, and relative ease of implementation. Adaptive PCA, also referred to as online PCA, utilizes various techniques to update principal components in response to new data, avoiding the need to process the entire data set. Oja's algorithm [29], which is based on neural networks, and Candid Covariance-Free Incremental PCA (CCIPCA) [30], which avoids covariance computation, provide efficient updates for streaming data. Recursive PCA [31] and

Matrix Stochastic Gradient (MSG) [32] offer recursive and matrix-based updates while maintaining computational efficiency. Incremental Singular Value Decomposition (SVD) constantly adjusts singular values and vectors, making it well-suited for real-time applications requiring an updated SVD. Every approach has unique advantages suited to particular data and application requirements. resilience of the acquired projection [27]. William Briguglio proposes a federated version of supervised PCA with its dual and kernel modifications, referred to as FeS-PCA, dual FeS-PCA, and FeSK-PCA, respectively [33]. Research on attacks against principal component analysis (PCA) has shown that disclosing eigenvectors can result in membership inference. Furthermore, when combined with knowledge of the data distribution, it can lead to data reconstruction attacks.

A. PROBLEM STATEMENT

There are so many challenges that affect the efficiency of the communication process. Background noise, weird voices, and occasionally missing sections of what is spoken due to signal problems can all obstruct communication. These issues arise in a variety of settings, such as when it is noisy outside or while talking on the computer. However, there is a solution known as speech enhancement. Our approach employs sophisticated computer methods and advanced technology to solve these issues. Its goal is to ensure we can hear each other clearly regardless of the noises or challenges surrounding us. Speech enhancement is a bridge that connects the difficulties we face today to a future where we can speak and share more effectively. The goal of this research is to design a technique that can eliminate noise and predict missing speech components while avoiding over processing, which can cause unnaturalness.

B. OUR CONTRIBUTION

This research article presents a novel voice enhancement method that combines stochastic approaches with online principal component analysis (PCA). The method is based on a stochastic residual approach. We construct and compare LMS based residual dominant component analysis (LRDCA) and APA based residual dominant component analysis (ARDCA) based on their efficiency in removing noise and sparsity. Stochastic approaches calculate the residuals of the sequentially arriving data. We collect the residuals in a matrix and update it after each iteration. Eigenvalues and eigenvectors of the residual matrix are calculated. Noise-free and predicted output is reconstructed. ARDCA not only removes noise and sparsity but also preserves crucial data features like dynamic range, energy, covariance, and RMS amplitude.

Our contribution mostly involves determining the eigen space of sequentially arriving data to ensure the highest possible reduction of noise and removal of sparsity while minimizing signal distortion.

II. AFFINE PROJECTION ALGORITHM

The Affine Projection Algorithm (APA) stands out among the stochastic approaches as the go-to option for various applications. APA is preferred due to its higher capacity to adjust to dynamic signal characteristics and successfully manage time-varying systems. Unlike LMS, which has a slower convergence rate, APA strikes a balance between convergence speed and computing complexity. It performs well when the signal statistics change or the system coefficients are sparse. APA is the method of choice for many adaptive signal processing jobs due to its ability to track quickly changing surroundings and its tolerance to noise. APA solidifies its place as the most popular methodology for computing residuals in various applications. Even when data sparsity is a significant obstacle, the Affine Projection Algorithm (APA) displays impressive performance. Its adaptability goes beyond just overseeing changing signal characteristics; it also performs well in situations where lots of data is missing. Applications with sparse data benefit greatly from APA's design, which efficiently captures and utilizes the minimal available information. Due to its adaptability, it can extract useful information from sparse datasets, enabling it to perform well in situations where more traditional algorithms could struggle. It can iteratively alter filter coefficients, making it incredibly efficient when data or system properties change over time [34]. APA's adaptability to different noise [35], [36] environments makes it a powerful tool for noise removal, providing enhanced performance and reliability across a wide range of signal processing applications.

Let

- X be the input data set.
- $S(h)$ be the input data at instance h.
- $T(h)$ be the desired or target output at instance h.
- $W(h)$ be the filter coefficients at instance h.
- $E_{priori}(h)$ be the prior error at instance h.
- $E_{posteriori}(h)$ be the posterior error at instance h.

Since we are dealing with adaptive environment and streaming data let us assume that we are keeping last $L+1$ input vectors, as shown in the equation at the bottom of the next page. The predicted output is given by

$$Y_{pred}(h) = S(h)W(h) = \begin{bmatrix} y_{pred,0}(h) \\ y_{pred,1}(h) \\ \vdots \\ y_{pred,L}(h) \end{bmatrix}$$

The desired or target output is given by

$$T(h) = \begin{bmatrix} t(h) \\ t(h-1) \\ \vdots \\ t(h-L) \end{bmatrix}$$

The priori error is evaluated as

$$E_{priori}(h) = \begin{bmatrix} t(h) - y_{pred,0}(h) \\ t(h-1) - y_{pred,1}(h) \\ \vdots \\ t(h-L) - y_{pred,L}(h) \end{bmatrix} = \begin{bmatrix} e_{priori,0}(h) \\ e_{priori,1}(h) \\ \vdots \\ e_{priori,L}(h) \end{bmatrix}$$

The weight vectors are updated as

$$W(h+1) = W(h) + \alpha S(h)[S(h)'S(h) + \beta I]^{-1} E_{priori}(h) \tag{1}$$

After updating weight vectors, the updated output is evaluated as

$$Y_{up}(h) = S(h)W(h+1) = \begin{bmatrix} y_{up,0}(h) \\ y_{up,1}(h) \\ \vdots \\ y_{up,L}(h) \end{bmatrix}$$

The posteriori error is calculated as

$$E_{posteriori}(h) = \begin{bmatrix} t(h) - y_{up,0}(h) \\ t(h-1) - y_{up,1}(h) \\ \vdots \\ t(h-L) - y_{up,L}(h) \end{bmatrix} = \begin{bmatrix} e_{posteriori,0}(h) \\ e_{posteriori,1}(h) \\ \vdots \\ e_{posteriori,L}(h) \end{bmatrix} \tag{2}$$

III. PROPOSED SCHEME

In real-world environments, speech signals are easily corrupted by noise. Noises can be grouped into stationary noise (unchanging as a function of time) and non-stationary noise (changing when shifted in time). Examples of non-stationary noise include street noise, train noise, babble noise (voices of other speakers), and instrumental sounds. The relationship between speech and noise in the time domain can be expressed as

$$s(h) = \tilde{s}(h) + n(h)$$

$s(h)$ is generated by blending the narrow band signal(digitized speech signal) and wide band signal $n(h)$ (noise). This speech enhancement scheme approximates the desired signal $\tilde{s}$ by minimizing the distortion factor as much as possible and predicting the missing values. The proposed stochastic residual residual based DCA comprises the following essential steps:

- We make use of a stochastic (either LMS or APA) adaptive filter to calculate posterior error, also referred as residuals in the context of signal processing.
- We initialize a column vector U of order m.
- The covariance of the residuals that are calculated in Equation 2 are assembled in C.
- At each iteration we make sure that the spectral norm of C does not exceed $2(\|X\|_F/l)$. If so calculate λ_{max} and corresponding eigen vector u of C at that particular iteration.
- U is updated by adding u next to the zero column. At the same time C is upgraded by subtracting $\lambda_{max} uu^t$.
- Once the complete data set is processed the resulting U contains all the Dominant Components.
- In case of LMS based residual DCA enhanced output is generated by

$$y_{enhanced} = UU^T s(h)$$

- For APA based residual DCA enhanced output is generated as

$$y = UU^T S(h)$$
$$y_{enhanced} = mean(y)$$

A brief summary of the proposed scheme is given in algorithm 1. Figure 1 illustrates the flowchart of the proposed methodology, providing a visual representation of the sequential steps involved in the process.

Theorem 1: The matrix C is symmetric and positive semi-definite.

Proof: Initially, C is a matrix with all entries zero, which means C is symmetric and positive semi-definite.

Next, C is updated by the addition of another symmetric matrix rr^T. Since the sum of two symmetric matrices is symmetric, the updated matrix $C + rr^T$ is also symmetric.

Now, let S be a non-zero matrix. We want to show that $C + rr^T$ is positive semi-definite. Consider the matrix product:

$$S^T(C + rr^T)S = S^T CS + S^T(rr^T)S$$

Since C is positive semi-definite, for any matrix S, the matrix $S^T CS$ is also positive semi-definite. This follows because:

$$S^T CS \geq 0$$

in the sense that all eigenvalues of $S^T CS$ are non-negative.

$$S(h) = \begin{bmatrix} s(h) & s(h-1) & \dots & s(h-L+1) & s(h-L) \\ s(h-1) & s(h-2) & \dots & s(h-L) & s(h-L-1) \\ \vdots & & & \vdots & \vdots \\ s(h-N) & s(h-N-1) & \dots & s(h-L-N+1) & s(h-L-N) \end{bmatrix}$$
$$S(h) = \begin{bmatrix} \mathbf{s}(h) & \mathbf{s}(h-1) & \dots & \mathbf{s}(h-L) \end{bmatrix}$$

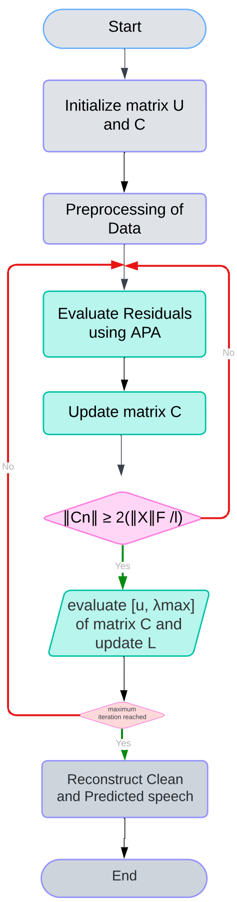


FIGURE 1. Flow chart of the proposed methodology.

For the second term, we compute:

$$S^T(rr^T)S = (r^TS)^T(r^TS)$$

Let $Y = r^TS$ be another matrix. Then we have:

$$S^T(rr^T)S = Y^TY$$

Since Y^TY is a Gram matrix, it is positive semi-definite, meaning:

$$Y^TY \geq 0$$

Thus, the sum of two positive semi-definite matrices, S^TCS and Y^TY, is positive semi-definite:

$$S^T(C + rr^T)S \geq 0$$

Therefore, $C + rr^T$ is positive semi-definite for any matrix S. □

Theorem 2:

$$\lambda_{max} \leq 2(\|X\|_F)^2/l$$

Proof: We have to evaluate the upper bound of the top eigenvalue of matrix C.

In the proposed algorithm, C is updated in two different ways. Firstly by the addition of covariance of residuals rr^T which causes increment in the top eigenvalue of C. But

---

Algorithm 1 APA Based Residual DCA Algorithm

Input: Given a data matrix X.

Initialization:Initialize an array U of order $m \times 1$ and an orthonormal matrix $V \in R^{p\times n}$. and a matrix C of order $m\times m$. Evaluate $\|X\|_F$

Prediction: For adaptive data projections weights are calculated recursively using 1 and residuals r (posteriori error) are calculated using 2. Detail is provided in the theory.

Update step:

1) The matrix C is updated as $C = C + r$
2) if
$$\|C_n\| \geq 2(\|X\|_F/l) \tag{3}$$
evaluate $[u, \lambda_{max}]$ of matrix C at that particular iteration.
3) U is updated by adding u next to its zero columns.
4) C is updated as
$$C = C - \lambda_{max}u_hu_h^T$$
5) For $n = n + 1$
U contains all the principal components.

Output:

$$y_h = u_hu_h^TS_h$$

---

this update occurs only when the loop condition given in Equation 3 fails, which means

$$\|C + rr^T\| \leq 2(\|X\|_F)^2/l \tag{4}$$

at the initial phase of each iteration. Secondly C is updated by the subtraction of $\lambda_{max}uu^t$ which causes reduction in the norm of C and hence decreasing its top eigenvalue.

The above discussion provides an upper bound for the top eigenvalue of C. □

Theorem 3:

$$\lambda_{max} \geq \|X(\|_F)^2/l$$

Proof: We have to evaluate the lower bound of the top eigenvalue of matrix C. Inside the loop

$$\|C + rr^T\| \geq 2(\|X\|_F)^2/l \tag{5}$$

when the above-given condition is true, we have

$$\|C\| \geq \|C + rr^T\| - \|rr^T\|$$

Using equation 5, we get

$$\|C\| \geq 2(\|X\|_F)^2/l - \|rr^T\|$$

We know that

$$r_h = (I - UU^T)S_h$$

where $(I - UU')$ is the matrix of projection and

$$\|r_h\|^2 \leq \|S_h\|^2 \leq (\|X\|_F)^2/l$$
$$\|r_hr_h^T\| \leq (\|X\|_F)^2/l$$

□

TABLE 1. Attributes of input and output signals.

Features	Input	LMS based DCA	APA based DCA
Energy	0.019070702636728	1.722074493484236e-04	0.002618372754349
Covariance	0.018942008303577	1.818069461057827e-04	0.003048488969921
Dynamic Range(dB)	1.192318190529117e+02	99.084867853758070	1.088359658601358e+02
RMS amplitude	0.137634074994162	0.013484598436184	0.055225555646827

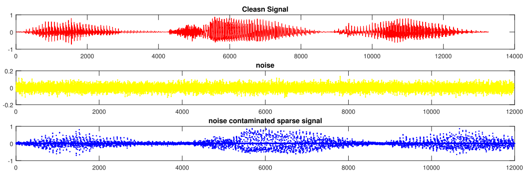


FIGURE 2. Signal Generation.

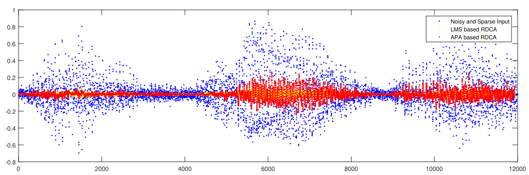


FIGURE 3. Noisy and Sparse Input Speech Signals and Enhanced Predicted Output using LMS basaed RDCA and APA based RDCA.

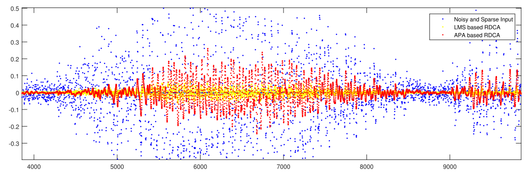


FIGURE 4. Zoomed-in view of the input and predicted enhanced output, highlighting detailed signal variations and improvements in noise reduction.

IV. EXPERIMENTAL RESULTS

To evaluate the effectiveness of the proposed algorithm, we use speech signals as input. The speech signals used are ''one two three''. Synthetic noise was added to the original speech signals by creating white Gaussian noise with a Signal-to-Noise Ratio (SNR) of 30 dB. The noise variance was computed using $sigma = 10^{-SNR/10}$, and the noise was thereafter generated based on this variance. The noise was subsequently included in the original signal to generate a noisy version, which was then employed to evaluate the speech enhancement method. This process ensured a consistent, noisy environment for evaluating the scheme's effectiveness. Following the addition of noise, the signal was made sparse by randomly choosing samples and setting to zero. The number of entries forced to be zeros depends on the sparsity factor. The term sparsity factor(η) is the ratio between the number of zero-valued entries and total entries.

$$\eta = \frac{no.\ of\ \ zero\ enries}{total\ number\ of\ \ \ entries}$$

The sparsity factor should be significantly greater than 0.5 for a matrix to be considered sparse. This was accomplished through the use of a random permutation to select the indices of the samples that will be assigned a value of zero. This stage guaranteed that the signal displayed sparsity, which is essential for evaluating the resilience of the speech enhancement algorithm in scenarios where the signal is sparse. Figure 2 shows the graphical representation of noisy and sparse input signals. The graphs in Figure 3 depict the efficacy of the APA-based RDCA in eliminating noise in comparison to the LMS-based RDCA. It illustrates a notable decrease in noise for both methods, but, the APA-based RDCA continually generates a more precise depiction of the original signal. This is further highlighted in zoomed in graph 4, where the APA-based RDCA exhibits improved noise reduction, leading to a more smooth waveform.

To verify the removal of sparsity, we calculate the number of zero entries in the input and output signals. The input contains ''5488'' zeroes, whereas the Output contains only ''65'' zeroes, hence verifying the removal of sparsity.

A comparison of crucial data attributes of input and outputs of both methods is given in Table 1. Energy measures a spoken signal's power or loudness over time. It affects speech volume and clarity, making it essential for speech enhancement. High signal energy improves speech audibility and understanding. However, severe energy reduction during noise suppression may degrade speech quality. Covariance examines the relationship between speech signal samples, which have similar amplitudes. This statistical measure reveals the correlation between speech signal samples. Speech enhancement algorithms must preserve dynamic range (DR) to maintain natural loudness when comparing quiet and loud areas. The RMS amplitude more accurately depicts loudness by representing the average magnitude of the waveform. Effective RMS level regulation balances clarity and distortion, ensuring excellent speech quality without noise. Table 1 clearly shows that LMS based DCA fails to maintain the integrity of input signals. The APA based RDCA is highly efficient in reducing noise and also excels in preserving the dynamic range and important properties of the speech signal. This makes it the superior technique for retaining signal integrity.

V. CONCLUSION

The experimental results show that the APA based RDCA method is superior to the LMS based RDCA in improving speech signals. It consistently reduces noise and sparsity while maintaining the overall quality and clarity of the speech. The RDCA method based on APA provides a more precise representation of the original signal while having minimal effects on important data attributes such as energy, covariance, dynamic range, and RMS amplitude. These findings highlight the significant potential of the applied algorithm in enhancing speech signals and its relevance in applications such as telecommunications, hearing aids, and voice-controlled systems.

DECLARATIONS

Conflict of interest The authors declare that they have no conflict of interest.

REFERENCES

[1] P. C. Loizou, *Speech Enhancement: Theory and Practice*. Boca Raton, FL, USA: CRC Press, 2007.

[2] T. Bahraini and A. N. Sadigh, "Proposing a robust RLS based subband adaptive filtering for audio noise cancellation," *Appl. Acoust.*, vol. 216, Jan. 2024, Art. no. 109755.

[3] J. Rownicka, P. Bell, and S. Renals, "Multi-scale octave convolutions for robust speech recognition," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 7019–7023.

[4] K. Kinoshita, T. Ochiai, M. Delcroix, and T. Nakatani, "Improving noise robust automatic speech recognition with single-channel time-domain enhancement network," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 7009–7013.

[5] Z. Xu, S. Elshamy, and T. Fingscheidt, "Using separate losses for speech and noise in mask-based speech enhancement," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 7519–7523.

[6] Z. Zhang, J. Geiger, J. Pohjalainen, A. E.-D. Mousa, W. Jin, and B. Schuller, "Deep learning for environmentally robust speech recognition: An overview of recent developments," *ACM Trans. Intell. Syst. Technol.*, vol. 9, no. 5, pp. 1–28, Sep. 2018.

[7] D. Michelsanti, Z.-H. Tan, S.-X. Zhang, Y. Xu, M. Yu, D. Yu, and J. Jensen, "An overview of deep-learning-based audio-visual speech enhancement and separation," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 29, pp. 1368–1396, 2021.

[8] N. Das, S. Chakraborty, J. Chaki, N. Padhy, and N. Dey, "Fundamentals, present and future perspectives of speech enhancement," *Int. J. Speech Technol.*, vol. 24, no. 4, pp. 883–901, Dec. 2021.

[9] J. Lin, A. J. de Lind van Wijngaarden, K.-C. Wang, and M. C. Smith, "Speech enhancement using multi-stage self-attentive temporal convolutional networks," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 29, pp. 3440–3450, 2021.

[10] A. Pandey and D. Wang, "Dense CNN with self-attention for time-domain speech enhancement," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 29, pp. 1270–1279, 2021.

[11] K. Tan and D. Wang, "Towards model compression for deep learning based speech enhancement," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 29, pp. 1785–1794, 2021.

[12] J. Zhang, M. D. Plumbley, and W. Wang, "Weighted magnitude-phase loss for speech dereverberation," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2021, pp. 5794–5798.

[13] V. Poluboina, A. Pulikala, and A. N. P. Muthu, "An improved noise reduction technique for enhancing the intelligibility of sinewave vocoded speech: Implication in cochlear implants," *IEEE Access*, vol. 11, pp. 787–796, 2023.

[14] S. Pradeep Kumar, A. Daripelly, S. M. Rampelli, S. K. R. Nagireddy, A. Badishe, and A. Attanthi, "Noise reduction algorithm for speech enhancement," in *Proc. Int. Conf. Signal Process., Comput., Electron., Power Telecommun. (IConSCEPT)*, May 2023, pp. 1–5.

[15] S. Bharti, P. Jha, M. Arora, and A. Kumar, "Speech enhancement and noise reduction in forensic applications," in *Proc. 26th Conf. Oriental COCOSDA Int. Committee Co-Ordination Standardisation Speech Databases Assessment Techn. (O-COCOSDA)*, Dec. 2023, pp. 1–5.

[16] N. Abajaddi, Y. Elfahm, B. Mounir, and A. Farchi, "A robust speech enhancement method in noisy environments," *Int. J. Electr. Comput. Eng. Syst.*, vol. 14, no. 9, pp. 973–983, Nov. 2023.

[17] B. Saha, Vasundhara, and R. V. Devi, "Siamese hybrid RNN model for speech enhancement: A novel approach for noise reduction in speech signals," in *Proc. Int. Conf. Circuit Power Comput. Technol. (ICCPCT)*, Aug. 2023, pp. 1676–1680.

[18] R. K. Anna and S. C. D. Uma, "Speech enhancement for noisy signals using adaptive algorithms," *i-Manager's J. Electron. Eng.*, vol. 13, no. 4, p. 26, 2023.

[19] S. Chhetri, M. S. Joshi, C. V. Mahamuni, R. N. Sangeetha, and T. Roy, "Speech enhancement: A survey of approaches and applications," in *Proc. 2nd Int. Conf. Edge Comput. Appl. (ICECAA)*, Jul. 2023, pp. 848–856.

[20] B. Colelough and A. Zheng, "Effects of dataset sampling rate for noise cancellation through deep learning," 2024, *arXiv:2405.20884*.

[21] B. Farhang-Boroujeny, *Adaptive Filters: Theory and Applications*. Chichester, U.K.: Wiley, 2013.

[22] X. Xie, N. Qian, and P. Qiu, "Online monitoring of air quality using PCA-based sequential learning," *Ann. Appl. Statist.*, vol. 18, no. 1, pp. 600–625, Mar. 2024.

[23] B. Liu and M. Lai, "Advanced machine learning for financial markets: A PCA-GRU-LSTM approach," *J. Knowl. Economy*, vol. 2024, pp. 1–35, Jun. 2024.

[24] S. K. Dash, S. Dash, S. Mahapatra, S. N. Mohanty, M. I. Khan, M. Medani, S. Abdullaev, and M. Gupta, "Enhancing DDoS attack detection in IoT using PCA," *Egyptian Informat. J.*, vol. 25, Mar. 2024, Art. no. 100450.

[25] J. Bian, D. Zhao, F. Nie, R. Wang, and X. Li, "Robust and sparse principal component analysis with adaptive loss minimization for feature selection," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3601–3614, Mar. 2022.

[26] C. D. Nguyen, C. H. Kim, and J.-M. Kim, "Gearbox fault identification model using an adaptive noise canceling technique, heterogeneous feature extraction, and distance ratio principal component analysis," *Sensors*, vol. 22, no. 11, p. 4091, May 2022.

[27] H. Vu, T. Tran, M.-C. Yue, and V. A. Nguyen, "Distributionally robust fair principal components via geodesic descents," 2022, *arXiv:2202.03071*.

[28] N. Migenda, R. Möller, and W. Schenck, "Adaptive dimensionality reduction for neural network-based online principal component analysis," *PLoS ONE*, vol. 16, no. 3, Mar. 2021, Art. no. e0248896.

[29] E. Oja, "Simplified neuron model as a principal component analyzer," *J. Math. Biol.*, vol. 15, no. 3, pp. 267–273, Nov. 1982.

[30] J. Weng, Y. Zhang, and W.-S. Hwang, "Candid covariance-free incremental principal component analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 8, pp. 1034–1040, Aug. 2003.

[31] I. Portnoy, K. Melendez, H. Pinzon, and M. Sanjuan, "An improved weighted recursive PCA algorithm for adaptive fault detection," *Control Eng. Pract.*, vol. 50, pp. 69–83, May 2016.

[32] R. Arora, A. Cotter, and N. Srebro, "Stochastic optimization of PCA with capped MSG," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 26, 2013, pp. 1–11.

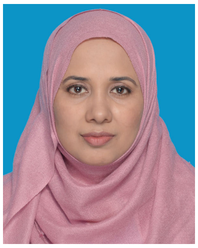
[33] W. Briguglio, W. A. Yousef, I. Traore, and M. Mamun, "Federated supervised principal component analysis," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 646–660, 2024.

[34] P. S. Diniz, *Adaptive Filtering*, vol. 4. Berlin, Germany: Springer, 1997.

[35] H. Zhao, F. Zhao, G. Wang, L. Zhou, and P. Song, "Affine projection algorithm for censored regression," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 68, no. 12, pp. 3602–3606, Dec. 2021.

[36] J. J. Jeong, "A robust affine projection algorithm against impulsive noise," *IEEE Signal Process. Lett.*, vol. 27, pp. 1530–1534, 2020.

RABIA SHARIF received the B.S. and M.S. degrees in mathematics from Lahore College for Women University, Lahore, Pakistan, where she is currently pursuing the Ph.D. degree, under the supervision of Dr. Shazia Javed. Her research interests include adaptive filtering, signal processing, and machine learning.



SHAZIA JAVED received the Ph.D. degree from University Sains Malaysia (USM), Penang, Malaysia, in 2014. She is currently an Assistant Professor in mathematics with Lahore College for Women University (LCWU), Lahore, Pakistan. Her research interests include optimization theory, numerical linear algebra, machine learning, and the application of adaptive filters in machine learning and signal processing.

MUSAED ALHUSSEIN received the B.S. degree in computer engineering from King Saud University, Riyadh, in 1988, and the M.S. and Ph.D. degrees in computer science and engineering from the University of South Florida, Tampa, FL, USA, in 1992 and 1997, respectively. Since 1997, he has been a Faculty Member with the Department of Computer Engineering, College of Computer and Information Science, King Saud University (KSU), Riyadh, Saudi Arabia, where he is currently a Professor. Recently, he has been successful in winning a research project in the area of AI for healthcare, which is funded by the Ministry of Education, Saudi Arabia. He is also the Founder and the Director of the Embedded Computing and Signal Processing Research (ECASP) Laboratory. His research interests include typical computer architecture and signal processing topics, with an emphasis on big data, machine/deep learning, VLSI testing and verification, embedded and pervasive computing, cyber-physical systems, mobile cloud computing, big data, eHealthcare, and body area networks.

UZMA BASHIR received the Ph.D. degree in mathematics from the School of Mathematical Sciences, Universiti Sains Malaysia, Penang, Malaysia, in 2014. She is currently an Assistant Professor with the Department of Mathematics, Lahore College for Women University, Lahore. She supervised many students in M.S. mathematics. Her area of research interests include computer-aided geometric design, differential and integral equations, and graph theory.

KHURSHEED AURANGZEB (Senior Member, IEEE) received the B.S. degree in computer engineering from the COMSATS Institute of Information Technology, Abbottabad, Pakistan, in 2006, the M.S. degree in electrical engineering (system on chip design) from Linkoping University, Sweden, in 2009, and the Ph.D. degree in electronics design from Mid Sweden University, Sweden, in June 2013. He is currently an Associate Professor with the Department of Computer Engineering, College of Computer and Information Sciences, King Saud University (KSU), Riyadh, Saudi Arabia. He has authored and co-authored more than 90 publications, including IEEE/ACM/Springer/Hindawi/MDPI journals and flagship conference papers. He has obtained more than 15 years of excellent experience as an Instructor and a Researcher of data analytics, machine/deep learning, signal processing, electronics circuits/systems, and embedded systems. He has been involved in many research projects as a principal investigator and a co-principal investigator. His research interests include the diverse fields of embedded systems, computer architecture, signal processing, wireless sensor networks, communication, and camera-based sensor networks, with an emphasis on big data and machine/deep learning with applications in smart grids, precision agriculture, and healthcare.

BHARAT BHUSHAN received the B.Tech. degree (Hons.) in computer science and engineering, the M.Tech. degree (Hons.) in information security, and the Ph.D. degree in computer science and engineering from the Birla Institute of Technology, Mesra, India, in 2012, 2015, and 2021, respectively. He was an Assistant Professor with the HMR Institute of Technology and Management, New Delhi, and a Network Engineer with HCL Infosystems Ltd., Noida. As a Research Fellow at Australian Institute for Machine Learning, he contributes to projects. He is currently an Assistant Professor with the Department of Computer Science and Engineering (CSE), School of Engineering and Technology, Sharda University, Greater Noida, India. In 2021 and 2022, he was with Stanford University, USA, listed in the top 2% scientists list. He earned numerous international certifications, such as CCNA, MCTS, MCITP, RHCE, and CCNP. He has published more than 150 research papers in various renowned international conferences and SCI-indexed journals, including *Journal of Network and Computer Applications* (Elsevier), *Wireless Networks* (Springer), *Wireless Personal Communications* (Springer), *Sustainable Cities and Society* (Elsevier), and *Transactions on Emerging Telecommunications* (Wiley). He has contributed with more than 30 book chapters in various books and has edited 20 books from the most famed publishers like Elsevier, Springer, Wiley, IOP Press, IGI Global, and CRC Press. He is a Series Editor of two prestigious Scopus-indexed book series named *Computational Methods for Industrial Applications* and *Future Generation Information System* (CRC Press and Taylor & Francis, USA). He has served as a keynote speaker (resource person) for numerous reputed faculty development programs and international conferences held in different countries, including India, Iraq, Morocco, China, Belgium, and Bangladesh. He has served as a reviewer/editorial board member for several reputed international journals. He is also a member of numerous renowned bodies, including IAENG, CSTA, SCIEI, IAE, and UACEE.