Received: 23 May 2020 | Revised: 1 June 2020 | Accepted: 2 June 2020
DOI: 10.37917/ijeee.16.1.14 Vol. 16 | Issue 1 | June 2020

Open Access

Iraqi Journal for Electrical and Electronic Engineering
Original Article

Enhancing Linear Independent Component Analysis: Comparison of Various Metaheuristic Methods

Nidaa Abdul Mohsin Abbas[1*], Hussein Mohammed Salman
[1] Faculty of IT, University of Babylon, Iraq
drnidaa_muhsin@ieee.org
[2] Faculty of Material Engineering, University of Babylon, Iraq
hus12ms@gmail.com

Correspondence
Nidaa Abdul Mohsin Abbas
Faculty of IT, University of Babylon, Iraq
Email: drnidaa_muhsin@ieee.org

Abstract
Various methods have been exploited in the blind source separation problems, especially in cocktail party problems. The most commonly used method is the independent component analysis (ICA). Many linear and nonlinear ICA methods, such as the radial basis functions (RBF) and self-organizing map (SOM) methods utilise neural networks and genetic algorithms as optimisation methods. For the contrast function, most of the traditional methods, especially the neural networks, use the gradient descent as an objective function for the ICA method. Most of these methods trap in local minima and consume numerous computation requirements. Three metaheuristic optimisation methods, namely particle, quantum particle, and glowworm swarm optimisation methods are introduced in this study to enhance the existing ICA methods.
The proposed methods exhibit better results in separation than those in the traditional methods according to the following separation quality measurements: signal-to-noise ratio, signal-to-interference ratio, log-likelihood ratio, perceptual evaluation speech quality and computation time. These methods effectively achieved an independent identical distribution condition when the sampling frequency of the signals is 8 kHz.

KEYWORDS: blind source separation, glowworm swarm optimisation, particle swarm optimisation, quantum particle swarm optimisation

I. Introduction

Blind source (Signal) separation (BSS), a popular signal treatment method has been used since the 1990s as an output of reaction amongst the neural networks, statistical processes, and information theory. BSS has become an excellent topic for many types of research and used in numerous applications, including the medical and medical sciences, signal telecommunications processing, sound and speech processing and image processing [1].
The following mathematical representation is used to model the process to understand BSS. Let the observation signal $x(t) = [x_1,x_2,\dots,x_n]^T$ represent n × 1 mixed-signal vector. In this mathematical representation, n denotes the number of sensors, t represents time coefficient and T means the transpose of the vector x. Each variable in the vector x represents the mixed signal recorded in the multiple-input/multiple-output format. This mixture system is shown as Equation (1):

$$x(t) = As(t) \qquad (1)$$

where $s(t)=[s_1,s_2,\dots,s_n]^T$ is a n × 1 latent vector with random independent and zero-mean non-normality distribution elements s_i , and A is an unknown n × n non-singular and full-rank mixing matrix. The model in Equation (1) represents the general linear model of the BSS and/or the independent component analysis (ICA) method [2], [3].
The statistical estimation approaches have been employed for this purpose. The main output of the BSS method expression is shown as Equation (2):

$$y(t) = Wx(t) \approx s(t) \qquad (2)$$

$y(t) = [y_1,y_2,\dots,y_n]^T$ represents the n × 1 separated vector and estimation of the real source signal, and W is an n × n estimated non-mixing matrix (i.e. separated matrix). The model is shown in Equation (2) represents the linear representation model of the separation process in BSS and/or ICA [4].
The core of the BSS methods to estimate and approximate the separated matrix W, which requires many pre-processes stages, such as whitening, data reduction, filtering and denoising [2, 4, 5].
Numerous methods are used to solve the BSS problem: ICA, non-negative matrix factorisation and sparse component analysis methods. Amongst the three methods, ICA is the most popular. ICA is an analysis method that depends on the

statistical properties of the mixed sources for decomposing the independent components [2, 5].

The core of the ICA algorithms mainly depends on two elements: the optimisation algorithm and the objective function. The objective function is used to maximise or minimise the mutual information amongst the components to be separated. The statistics of the ICA method, including the robustness and consistency, are based on the objective function. The algorithm used for enhancing the performance of the ICA is an optimisation method. The algorithmic properties of the ICA, which include the stability, convergence speed, and memory requirements, depending on the optimisation algorithm [2].

Most traditional ICA methods use artificial neural networks as optimisation methods. These methods are based on gradient functions as an objective function to estimate the components [2, 6]. The gradient functions suffer from trapping in the local minima of the search space and time consumption in the training and learning processes. Although the neural networks based on ICA methods can use the informatics theory functions as the entropy, these networks must still perform training and learning processes and may be trapped in local minima [7, 8].

The BSS problem can also be solved by using metaheuristic optimisation methods, such as a genetic algorithm (GA) [9], particle swarm optimisation (PSO) [10] and simulated annealing (SA) [11] algorithms. The ICA methods use the information-theoretic concepts, such as differential entropy, negentropy, mutual information, and maximum likelihood, as objective functions [7].

Two new metaheuristic optimisation methods are proposed in this study to enhance the performance of the ICA linear mixture: quantum PSO (QPSO) [12] and glowworm swarm optimisation (GSO) [13]. These proposed optimisation methods employ the following three objective functions: negentropy, differential entropy, and MI.

II. Related Work

The existing literature that introduced methods as solutions for the ICA methods based on metaheuristics is explored in this section.

Liu et al. (2006) introduced a hybrid method of PSO and a neural network to solve a BSS problem. Their method used a feed-forward neural network optimised by PSO and computation time as the evaluation metric [14].

Song et al. (2007) introduced a method to enhance the performance of the post-nonlinear (PNL) for the nonlinear mixture. They also used PSO to determine the inverse nonlinear function and perform approximation via high-order odd polynomial (HOOP) function. The proposed method used mutual information as a fitness function in PSO and the natural gradient to compute the separated matrix. Furthermore, their method utilised source to distortion ratio measurement for the evaluation [15].

Yu et al. (2007) presented a method to improve PSO by using 'local deep search' and 'migration operation' to avoid trapping in local minima. The improved PSO was then used to estimate the parameters of the polynomial $H(y_i)$. They utilised the mutual information as the evaluation function and a parameter in the natural gradient algorithm to estimate the linear separated matrix in the ICA algorithm. The transposition error parameter was also employed as the evaluation metric in their method [16].

Cai and Tian (2011) used the HOOP to set the inverse nonlinear function, wherein the parameters were optimised by PSO. They utilised the joint entropy estimation to convert mutual information. The separated signals may become unsatisfactory due to the unstable mutual information in the finite sample length. They also used the Gaussian mixture model (GMM) to set the pdf of separated sources. The correlation coefficient measurement is used as the evaluation metric [10].

Kurihara and Jin'no (2013) proposed a method for nonlinear ICA by using the radial basis function (RBF) network. They employed PSO to optimise the parameters of the RBF. Moreover, they used the Gaussian distribution function as the cost function in RBF. Furthermore, their proposed method can use other functions, such as linear, cubic, thin plate, multi-quadratic or inverse multi-quadratic functions. Their research used the mean square error as an evaluation metric [17].

Lee and Koehler (1997) presented a method to solve the nonlinear ICA by using high order polynomial. Their method assumed that nonlinearity can be approximated by polynomials of the nth order. Moreover, their proposed algorithm focused on parametric sigmoidal nonlinearity and high-order polynomials and used a signal-to-noise ratio (SNR) as an evaluation metric [18].

Ziehe et al. (2001) introduced a method for nonlinear ICA by using the temporal decorrelation separation (TDSEP) and alternating conditional expectation (ACE) algorithms. The TDSEP algorithm was used for the separation. Their method employed the ACE algorithm to map the nonlinearity from the observations and the TDSEP algorithm to separate the sources. Their method used the correlation coefficient as an evaluation metric for the separation process [19].

Xiong and Huang (2001) proposed an algorithm for ICA based on power series. The proposed algorithm used the gradient ascent, mutual information and traditional linear approximation to nonlinear problems by using Taylor expansion. For the evaluation, their method used the convergence speed parameter as the evaluation metric [20].

Tan and Wang (2001) introduced an algorithm for BSS by exploiting the GA to minimise the cost function in the implementation of the system through a neural network. Moreover, they used the high-order statistic as an objective function to maximise the independence of the mutual information parameter. Their research employed the cross-correlation measurement as the evaluation metric [21].

Puntonet et al. (2002) exploited SA with competitive learning manners in neural networks to solve the BSS problem. Their research indicated that the root means square error (RMSE) was used as an evaluation measurement [11].

Eriksson and Koivunen (2002) presented an algorithm for BSS by using an additive theorem for nonlinear mixing. They presented a table of the mixing operators and their functions. These functions belong to nonlinear instantaneous ICA models (i.e. addition theorem) and continuous and strictly monotonic functions (e.g. cx, ex, tan(x) and tanh(x)).

Their research used the signal-to-interference ratio (SIR) as an evaluation metric [22].

Rojas et al. (2002) presented a method for the nonlinearity of ICA based on the GA and the approximation of mutual information function as an objective function of their method. They also employed the benefits of GA to explore and separate the mixed signals. Their method combined GA with the natural gradient descent and used the plotting of the signals as an evaluation metric [9].

Sole et al. (2002) introduced a method that depends on minimising mutuality information between mixed signals. Their method used the polynomial parameterisation model. The one-hidden layer multilayer perceptron (MLP) parameterisation model was then used. Moreover, the gradient descent algorithm was applied as the cost function of the network, and the polynomial was used to estimate the inverse of the nonlinear function instead of the sigmoidal functions. The evaluation metrics used in their research included the SNR evaluation measurement and the linear predictive coefficients (LPC) of the sources, and signals were separated to depict these metrics [23].

Almeida (2003) proposed an algorithm for nonlinear mixture in the ICA (i.e. 'MISEP'). MISEP is an extension of the INFOMAX algorithm used for linear separation. The entropy function was used in their research as the objective function, wherein the Jacobian matrix is a parameter of this function. The one-hidden layer MLP neural network was also employed to compute the Jacobian matrix. The gradient descent was used to update the weights of the MLP network with an odd polynomial function. The SNR measurement was used as an evaluation metric [24].

Rojas et al. (2004) presented an algorithm of the ICA based on the neural network with competitive learning. Their algorithm used SA for the random generation of weights. The GA was then utilised to update these weights. They used mutual information as an objective function in the GA. Their proposed algorithm could be used in linear and nonlinear ICA. Their research also used the crosstalk parameter as a metric of the evaluation process [25].

Karvanen and Tanaka (2004) proposed an algorithm that applies the PNL mixture method, which used the Pearson algorithm in the ICA method with the temporal decorrelation manner. Their method utilised the Gaussian to approximate the observation signals before the transformation in the nonlinearity formula. Their research used SIR as an evaluation metric [26].

Dias et al. (2009) proposed a method to solve the BSS of PNL mixture using evolutionary computation and Gaussianization. This method used the immune system, wherein the Gaussianization was utilised. Their method combined the FastICA algorithm as a local search mechanism with the opt-aiNet as an evolutionary global search method. They used negentropy as the measurement of the Gaussianity. The inverse nonlinear transformation was set as a fifth-order polynomial with odd power. Their research used several parameters (i.e. convergence speed (number of iterations), elapsed time in each loop (ms) and convergence time (min)) as evaluation metrics [27].

Oveisi et al. (2012) introduced a method to analyse the EEG data by using the nonlinear ICA. The EEG data were already mixed in their proposed method. Moreover, the method used the PNL method as a nonlinear mixture and GA for mutual information optimisation. Their research used the classification accuracy rate as an evaluation measurement [28].

Ehsandoust et al. (2016) proposed an algorithm for nonlinear ICA by using the Jacobian matrix as a nonlinear mixture and then it is inverse as the separation matrix. The Jacobian function was used as an inverse of the nonlinear function. They also utilised artificial signals of the sources (i.e. sine and sawtooth waves). Smoothing algorithms were employed as the reduction methods of the mixed data. Their research utilised the RMSE as an evaluation measurement [29].

III. BACKGROUND THEORIES

A. Independent Components Analysis (ICA)

The most popular and trusted algorithm used for BSS is ICA. It is a statistical algorithm based on the statistical computations of the observation data. The ICA of the vector x includes an estimation of a generic form of the observed signal, as shown in Equation (1).

In this form, the components s_i that appeared in the vector $s = (s_1, .., s_n)^T$ are assumed to be independent. The matrix A is an m × n fixed 'mixing' matrix. The noise vector in this model is slightly observed and thus can be aborted. The noiseless (noise-free) form considers a possible approximation of the realistic disturbed form [2].

The main processes in the ICA are the estimation and recovery of the original signals and the performance of the separation process. Linearly, later processes include an estimation of an inverse of the mixed matrix A, as shown in Fig. 1.

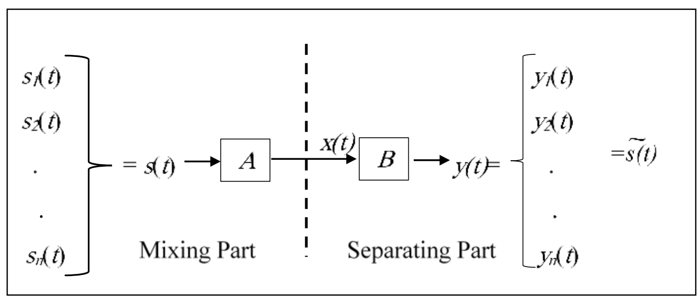

Fig. 1: Sketch of the linear ICA system.

This approach is used to estimate a separated matrix B and find the variables y_j as shown in Equation (3). The calculation of the s_j up to permutation and scaling is also shown in Equation (3).

$$y(t) = Bx(t) \tag{3}$$

Hence, B is an estimation of the inverse of A. In most cases, the sources s_j are found individually such that the y_j is considered to be an estimation of the sources s_j [2].

B. Measures of Non-Gaussianity in ICA

This subsection presents the measures of the Gaussianity and non-Gaussianity of the observation signals in BSS and ICA methods. These measures, which include kurtosis, negentropy, mutual information, and maximum likelihood, are discussed below.

- **Kurtosis**

Kurtosis is commonly used to measure Gaussianity and represents the fourth-order cumulant. This measure is defined as indicated in Equation (4) [2, 7]:

$$K(x) = E[x^4] - 3(E[x^2])^2 \quad (4)$$

For the unit variance of x, the kurtosis equation can be rewritten as $E\{x^4\}-3$. This mathematical expression proves that kurtosis is a normalised model of the fourth-order moment $E\{x^4\}$. The fourth moment is $3(E\{x^2\})^2$ for a Gaussian distribution. Therefore, the kurtosis value is zero for a Gaussian distribution. The kurtosis value of most non-Gaussian distributions is nonzero.

Kurtosis could be defined using signs (−, +, 0): (negative) sub-Gaussian, (positive) super-Gaussian and (zero) Gaussian. The kurtosis for an ICA method can be used as a theoretical optimisation criterion [8]. Kurtosis consumes lesser computation time than that of other measurements.

- **Negentropy (Negative Entropy)**

Negentropy depends on the amount of the theoretical information of the entropy. The entropy measures the randomness of the variable. The entropy for a discrete variable is shown as Equation (5):

$$H(x) = -\sum P(x)\log\ P(x) \quad (5)$$

where H denotes the entropy of the observation signals and an estimation of the original signals, p is the probability of P and x represents possible values of x. The entropy generally uses the logarithm with base two. The Gaussian distribution variables have larger entropy than that of other variables. The negentropy concept is used to measure the Gaussianity of the components, as shown in Equation (6) [2, 7]:

$$J(x) = H(x_G) - H(x) \quad (6)$$

where x_G denotes the Gaussian variable. If the variable is Gaussian, then the $J(x)$ is zero and generally non-negative. The negentropy is nonparametric and expensive in computations despite its robustness in statistical considerations.

Thus, negentropy is unfit for the computations. Estimating negentropy using the definition requires an approximation of the probability density function (PDF). Therefore, the negentropy is useful and commonly used in approximating the PDF, as shown in Equation (7) [2]:

$$J(x) \propto [E\{G(x)\} - E\{G(v)\}]^2 \quad (7)$$

where v and x denote the Gaussian vectors with zero means, and G represents the quadratic function.

Mutual Information

The mutual information amongst variables, x_i, i = 1,2,..,m, can be defined by using the differential entropy, as shown in Equation (8):

$$MI(x_1, x_2, ..., x_m) = \sum_{i=1}^{m} H(x_i) - H(x) \quad (8)$$

Mutual information is a metric used to measure the dependence of the variables. This information is equivalent to the Kullback Leibler divergence concept between the joint density $f(x)$ and the production of its marginal densities. When the x_i is independent, mutual information will be equal to zero due to the absence of information between the two.

Mutual information measures the information-theoretic of the independent variables. Thus, mutual information can be used as a measurement to estimate the ICA transformation. Using this method instead of an estimate of the model method, the ICA of vector x can be defined as the invertible model ($s_i = Wx_i$), where the estimation of the separation matrix W is used for minimising the mutual information amongst the components s_i [2].

C. Pre-Processing of the ICA

The ICA method has several pre-processing operations, such as filtering, denoising and data reduction. However, centering and whitening are the two most common operations. This subsection introduces these operations and their equations.

- **Centring**

This concept is strongly related to the central moment, which is the most important pre-processing operation for the ICA methods. Centering refers to the computation of the mean of the mixed-signal vectors, and this mean is subtracted from the mixed vector itself:

$$x' = x - E[x] \quad (9)$$

A result vector is called the mean or zero-mean vector. The mean vector is then added to the computed vector, as shown in Equation (10):

$$s = s' + A^{-1}E[x] \quad (10)$$

After this process, the mixing matrix A remains. This process does not affect the estimation of the mixing matrix. The reconstruction of the subtracted mean vector after the estimation of the components is performed by adding A^{-1} $E\{x'\}$ to the zero-mean components [2, 4].

- **Whitening**

The subsequent pre-processing stage is called the whitening process, which is also known as the sphering process. After this process, uncorrelated mixed signals x and the unit variance of the observed data are obtained. This process can be achieved by applying the model in Equation (11):

$$\tilde{x} = \Lambda D^{-1/2} \Lambda^T x \quad (11)$$

Columns represent the eigenvectors of , and the diagonal D represents the eigenvalues of $E[xx^T]$. The whitening process aims to make the mixing matrix orthogonal [2, 7].

D. Optimisation Methods

The optimisation algorithms are mostly applied in many areas such as economics, social sciences, and engineering. These methods are used to determine the best solution amongst various optimum solutions of a particular problem and then initialise scientific support to guide the decision-makers. Optimisation methods require two processes. The first process initialises the problem to be optimised. This process requires problem exploration using a suitable objective function, constraints and decision variables of the problem. The second process includes a numerical approach, wherein the solution of the problem is determined, the proposed solution is examined and the desired state is realised [12].

A metaheuristic approach refers to an iterate generation approach that adopts the secondary heuristic using a combination of the intelligent exploration concepts and the search area exploitation. The metaheuristic algorithms use the learning algorithms for information structuring to find an

efficient solution [12]. The exploitation idea in these algorithms represents the local search, whilst the exploration issue represents the global search.
Metaheuristic strategies are naturally stochastic and local random search algorithms. Swarm intelligence, SA and evolutionary algorithms are examples of metaheuristic approaches. This study focuses on swarm intelligence metaheuristic approaches, particularly PSO [30], QPSO [31] and GSO [13].

- **Particle Swarm Optimisation (PSO)**

Kennedy and Eberhart (1995) presented the PSO algorithm in their study [30]. PSO is one of the search approaches that use the heuristical population search. In this method, each search term called particle comprises two main parameters: position and velocity. Each particle searches the best position in the local search space called the local best position and stores all its positions in its memory. These positions are defined as the experience of the current particle in the current dimension. The particle swarm discovers new positions in other dimensions through the search process, and these new positions are called the new experience. The final and main task of the particle swarm is to find a global best position during the search iteration. The global best position represents the best position amongst the local best positions in n-dimensions of the current particle (12). The position and the velocity of each particle are respectively calculated in Equations (12) and (13) for n iterations:

$$v_i(t+1) = wv_i(t) + c_1 r_1(t)(pbest_i(t) - x_i(t)) + c_2 r_2(t)(gbest_i(t) - x_i(t)) \tag{12}$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \tag{13}$$

where v denotes a velocity of a particle P, x represents the position of P, *pbest* denotes the best local position of the current P, *gbest* denotes the best global position for all Ps in the search area in n-dimension and w denotes an inertia weight (for the convergence speed). Therefore, c_1 and c_2 are the acceleration parameter constants, and r_1 and r_2 denote the two parameters valued randomly in the range [0 to 1].

- **Quantum Particle Swarm Optimisation (QPSO)**

The QPSO approach is a revision of the PSO approach. QPSO does not have the velocity parameter; it only requires a small number of parameters and is simple in application [31]. This approach provides good performance in the solution of various problems [12]. This technique is described as follows.
QPSO assumes that each particle searches in an area with a δ potential on a certain dimension nearby the point pij. The particle swarm can generally be represented in a certain dimensional area, with a center p. The Schrödinger formula is used to solve the dimensional δ potential. Based on this formula, the PDF Q and the distribution function F can be defined as in Equations (14) and (15), respectively.

$$Q(X_{ij}(t+1)) = \frac{1}{L_{ij}(t)} e^{-2\left|p_{ij}(t) - x_{ij}(t+1)\right|} / L_{ij}(t) \tag{14}$$

$$F(X_{ij}(t+1)) = e^{-2\left|p_{ij}(t) - x_{ij}(t+1)\right|} / L_{ij}(t) \tag{15}$$

where $L_{ij}(t)$ is calculated using the Monte Carlo estimation approach to represent a standard deviation. The particle position can also be calculated as in Equation (16):

$$X_{ij}(t+1) = P_{ij}(t) \pm \frac{L_{ij}(t)}{2} \ln(1/u), \quad u = rand(0,1) \tag{16}$$

In the evaluation, the $L_{ij}(t)$ algorithm uses the mean best position m, which is a global point of the population, and the pbest of all particles.

$$m(t) = (m_1(t), m_2(t), \ldots, m_n(t)) = \left(\frac{1}{M}\sum_{i=1}^{M} P_{i,1}(t), \frac{1}{M}\sum_{i=1}^{M} P_{i,2}(t), \ldots, \frac{1}{M}\sum_{i=1}^{M} P_{i,n}(t) \right) \tag{17}$$

where M denotes the size of the population, and Pi represents the pbest of the particle i. The $L_{ij}(t)$ is calculated as shown in Equation (18):

$$L_{ij}(t) = 2\beta * \left| m_j(t) - X_{ij}(t) \right| \tag{18}$$

The position of the particle i is calculated as shown in Equation (19):

$$X_{ij}(t+1) = P_{ij}(t) \pm \beta * \left| m_j(t) - X_{ij}(t) \right| * \ln(1/u) \tag{19}$$

where β represents the contraction–expansion factor, which is the control parameter of the algorithm convergent [32].

- **Glowworm Swarm Optimisation (GSO)**

GSO is one of the swarm intelligent optimisation methods introduced in 2005 by Krishnanad and Ghose [33]. GSO describes the special-level conditions that facilitate the employment of an agent swarm in a signal space to realise sub-swarm convergence through multi-sources. The idea of GSO is taken from the glowworms (which are also known as lightning bugs or fireflies). According to the behaviours of the glowworm bugs, the glowworm swarm can change its luciferin intensity emanation at any moment. Based on this feature, the GSO method optimises the problem by approximating the intensities of the glows to the optimum fitness value of the function. Practically, the agent glows with low intensities (dark) collect around the agent glow with high intensities (bright).
The GSO method splits the agent glows into small sets and converges these glowworms around points with high values of the fitness function. Based on this advantage, the GSO algorithm can be used as an identification model of the multi-peaks of the multimodal function. This feature is exclusive to the GSO method; this method is the first to provide an indirect solution to solve the optimisation problems [13].
The GSO technique generally includes the following five key stages: update the luciferin stage, select the neighbourhood stage, movement of the probability-computer stage, movement stage, and updating the stage of the decision radius. The GSO algorithm is described as follows [33].

1. *Luciferin-Update Stage*. Updating of the luciferin stage depends on the fitness and previous luciferin values, and its rule is given in Equation (20):

$$l_i(t+1) = (1-p)l_i(t) + \gamma \text{Fitness}(x_i(t+1)) \tag{20}$$

where $l_i(t)$ represents the luciferin value of the agent glow i at time t, p is the luciferin decay constant, γ is the luciferin enhancement constant, $x_i(t+1) \in R^M$ is the location of glowworm i at time $t+1$ and Fitness($x_i(t+1)$) denotes the value of the fitness at the location of glowworm $i$at time t+ 1.

2. *Neighbourhood-Select Stage.* Neighbours (t) of agent glow i at time t comprises the brighter ones and can be solved as Equation (21):

$$N_i(t) = \{ j : (t) < r_i^d(t); l_i(t) < l_j(t)\} \qquad (21)$$

where d denotes a Euclidean distance between agent glow i and agent glow j at time t. $r_i^d(t)$ denotes a decision radius of glowworms i at time t. l_i denotes a luciferin of agent glow i, and l_j denotes the luciferin of glowworm j.

3. *Moving Probability-Computer Stage.* A glowworm uses the probability function to reposition the current glowworm into the other glowworms that are more luciferin level than the current glow luciferin level. The probability $P_{ij}(t)$ of glowworm i moving towards its neighbour j can be stated as follows:

$$P_{ij}(t) = \frac{l_j(t) - l_i(t)}{\sum_{k \in N_i(t)} l_k(t) - l_i(t)} \qquad (22)$$

4. *Movement Stage.* Suppose glowworm i selects an agent to glow $j \in N_i(t)$ with $P_{ij}(t)$, the discrete-time model of the movement of glowworm i is shown in Equation (23):

$$x_i(t+1) = x_i(t) + s\left(\frac{x_j(t) - x_i(t)}{\left\| x_j(t) - x_i(t) \right\|}\right) \qquad (23)$$

where $\|\cdot\|$ denotes the Euclidean norm parameter, s is the step-size, x_i represents the location of glowworm i and x_j represents the location of glowworm j.

5. *Decision Radius Update Stage.* In each update, the decision radius of glowworm i is given as follows:

$$r_i^d(t+1) = \min\{r_s, \max\{0, r_i^d(t) + \beta(n_t - |N_i(t)|)\}\} \qquad (24)$$

where β is a constant, r_s denotes the sensory radius of glowworm i and n_t is the control parameter of the number of the neighbouring agents.

IV Evaluation Measurements of the Source Separation

The evaluation measurements for the separation process performance can be categorised into two main groups: objective and subjective measurements.

1- Objective Measurements

These measurements are based on numerous physical aspects, including an acoustic factor or its conversion level, and several mathematical calculations.

The objective measurements of the sound quality can be computed from the sound with and without noise (original) by exploiting a certain mathematical model. These measurements do not require any examination of a human listener. Moreover, the measurements are relatively cheap, with low computations and less consumption of the time computation.

- **SNR Measurement**

The relationship between the source signal and the separated signal can be evaluated using various metrics, such as SNR and SIR. SNR is the most popular objective measurement used to evaluate speech quality [34]. Mathematically, this measurement requires a simple computation despite its assumption that the original and the distorted sounds are available.

The SNR measurement can be computed as shown in Equation (25):

$$SNR = 10\log_{10} \frac{\sum_{n=-\infty}^{\infty} s^2(n)}{\sum_{n=-\infty}^{\infty} (s(n) - \hat{s}(n))^2} (dB) \qquad (25)$$

where $s(n)$ represents a clean (original) signal data, and $\hat{s}(n)$ represents the separated signal. The normal scores of the SNR range from 0 to 1, in which the normal results are close to 0 and vice versa [34].

- **SIR Measurement**

The SIR measurement, which is similar to SNR, is used to evaluate the relationship between the source signal and the separated signal samples. The SIR can be calculated as shown in Equation (26) [34]:

$$SIR = 10\log_{10} \frac{\left| \sum_{n=-\infty}^{\infty} s(n)^2 \right|}{\left| \sum_{n=-\infty}^{\infty} \hat{s}(n)^2 \right|} (dB) \qquad (26)$$

where $s(n)$ denotes the clean (source) signal samples, and $\hat{s}(n)$ represents the separated signal.

- **Log-Likelihood Ratio (LLR) Measurement**

Various distance measurements use an estimation of LPC for the source (clean) signal samples and the distorted sound signal samples [34, 35, 36].

LLR measurement is one of the distance measurements that exploit the estimation of the LPC vectors of the original and the separated (distorted) sound signals. The LLR measurement can be computed as shown in Equation (27):

$$LLR(a_d, a_c) = \log\left(\frac{a_d R_c a_c^T}{a_c R_d a_c^T}\right) \qquad (27)$$

where $\boldsymbol{a}_c$ represents the LPC estimated vector of the source (original) sound signal, $\boldsymbol{a}_d$ represents the LPC estimated vector of the separated sound signal, $\boldsymbol{a}^T$ represents a transpose of $\boldsymbol{a}$, $\mathbf{R}_c$ represents an auto-correlation matrix of the original sound signal, c is the original sound signal and d is the separated sound signal.

2- Subjective Measurements

Plotting and playing of the signals are used for the subjective evaluation of the source (original) and the separated (distorted) sound signals. The separated signals and the accuracy in the results are observed by plotting the source (original), mixed (observation) and separated (distorted) signals in a graphical window and playing the audio of all signals (source, mixing, recovered).

- **Perceptual Evaluation Speech Quality (PESQ) Measurement**

The subjective measurements require human listeners, which makes it more expensive with more computational requirements than the objective measurements. However, some objective measurements can be employed to behave similarly to the subjective measurement and provide good estimation quality [34].

PESQ is the most common objective measurement used in evaluating the quality of speech signals in the BSS methods. The PESQ measurement compares the source (original) sound signals *s* with the separated (estimated) sound signals *y*. The main task of this measurement is to predicate the quality that could be achieved by the subjective listeners. The normal range scores of the PESQ-MOS measurement are between 1.0 (bad) and 4.5 (no distortion).

V. Proposed System

The proposed system comprises four phases. In the first phase, the raw data (i.e. source signals as sounds or speeches) are initialised under some assumptions to be suitable for the mixing process, and the mixing process of the sound and speech signals is performed in particular conditions. The mixed signals are then passed into the second phase after the mixing process. The second phase includes the execution of the ICA algorithm. The third phase represents the core of the proposed system, which includes the selection of the optimisation method. The optimisation methods (PSO, QPSO, and GSO) are used and the parameters are set to appropriate with the ICA. In this phase, the negentropy, entropy and mutual information functions for the linear mixture are suggested to be used as an objective (contrast) function. In the last phase, numerous objective evaluation measurements are used to evaluate the performance of the optimised ICA method. Fig. 2 shows the sketch of a proposed algorithm.

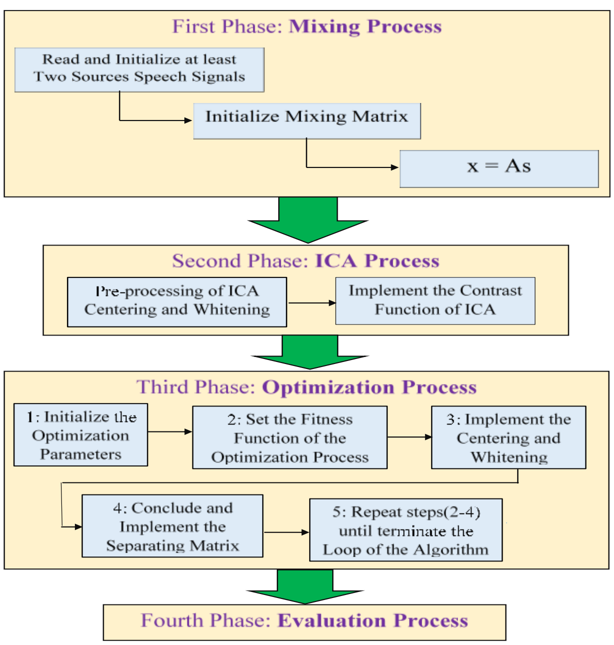

Fig. 2: General Block Diagram of the Proposed System.

- **Experimental Results and Discussion**

Table 1 shows a list of mixed sounds. The mixing matrix and its good conditions for 12 mixed cases of two source signals are formed from 17 sound files taken from the International Telecommunication Union (ITU) and the University of Dallas databases [37].

TABLE I
SELECTED FILES AND MIXING MATRIX

Mixed Case No.	Source Signal File Names	Length	Mixing Matrix	Well condition (Cond.)	Encode of Mixing Signal Figures	Frequency
1	source 11 source 22	50000	−0.6010 −0.6429 0.7899 −0.6822	1.1913	A12	8 KHz
2	Source 4 Source 7	50000	1.3675 −1.5102 −1.3945 −1.3716	1.0597	B12	
3	09m 15f	23323	−0.7115 −0.6815 0.6615 −0.4922	1.2782	C12	
4	Julia8 22m	21582	−0.7049 0.9661 −0.6066 −0.9998	1.0987	D12	
5	source 11 source 2	50000	−0.9514 −1.4418 1.7464 −1.7376	1.5068	E12	
6	source 22 source 2	50000	1.0133 −1.5314 −1.1377 −1.6548	1.4982	F12	
7	source 2 source 4	50000	−1.5299 −0.6929 0.8602 −0.6929	1.9188	G12	
8	source 4 source 9	50000	−1.0206 −1.3245 −0.9696 1.3740	1.3558	H12	
9	source 7 source 8	50000	1.3605 1.9283 −0.6674 1.5213	1.9750	I12	
10	Rich8 source 8	50000	1.9103 1.4365 1.7847 −1.5739	1.2273	J12	
11	Ray8 Rich8	61038	1.2319 −0.7851 1.4445 1.6247	1.5375	K12	
12	source 2 source 9	50000	−1.2848 1.8279 1.9384 1.3240	1.0525	L12	

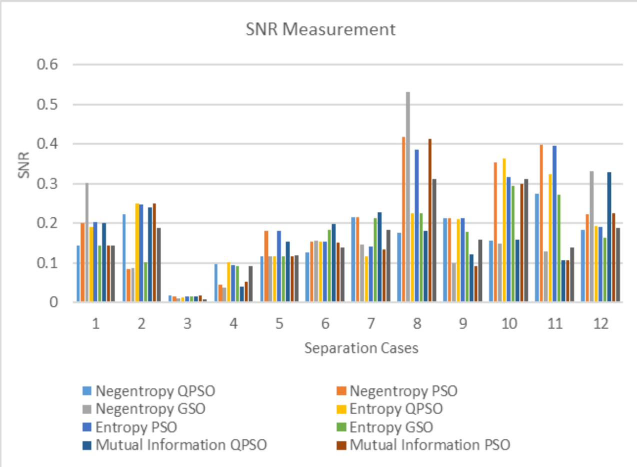

Fig. 3: SNR Measurement for the Proposed Algorithm

The Fig. 3 shows that the proposed algorithms exhibited good results with the negentropy, entropy and MI functions in six, four and five cases, respectively. According to the SNR measurement, the proposed algorithms based on the negentropy function are better than those of entropy and mutual information functions.

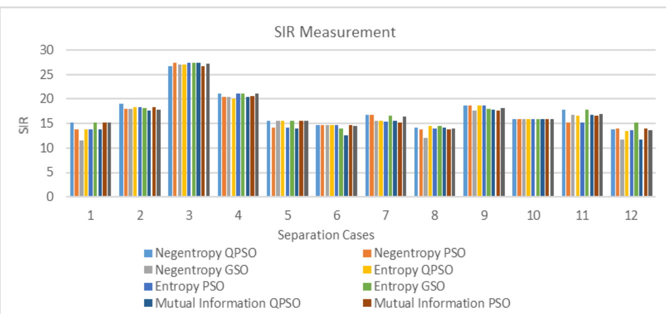

Fig. 4: SIR Measurement for the Proposed Algorithm

The Fig. 4 shows that the proposed algorithms exhibited good results with the negentropy, entropy and mutual information functions in nine, three and two cases, respectively. According to the SIR measurement, the proposed algorithms based on the negentropy function are better than those of entropy and mutual information functions.

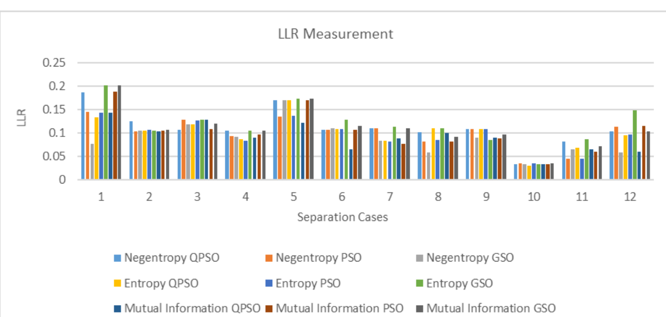

Fig. 5: LLR Measurement for the Proposed Algorithm

The Fig. 5 shows that the proposed algorithms exhibited good results with the negentropy, entropy and mutual information functions in five, two and four cases, respectively. According to the LLR measurement, the proposed algorithms based on negentropy function are better than those of entropy and mutual information functions.

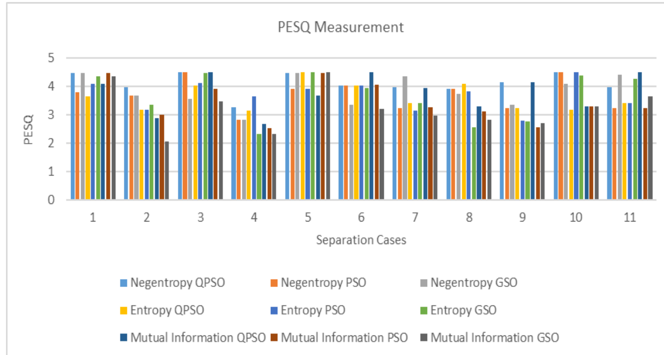

Fig. 6: PESQ Measurement for the Proposed Algorithm

Amongst the proposed algorithms, the PESQ measurement exhibited the most reliable results, which are nearest to the real playing of the separated sounds compared with that of the source sounds. Fig. 6 shows that the proposed algorithms exhibited good results with the negentropy, entropy and mutual information functions in five, four and three cases, respectively. According to the PESQ measurement, the proposed algorithms based on all the used functions demonstrated good results in the separation process. Most of these results were more than or equal to 4.

VI. CONCLUSION

Two metaheuristic optimisation methods (i.e. QPSO and GSO) are proposed in this study to enhance the performance of existing ICA methods. The proposed methods exhibited good results under various quality measurements, including SNR, SIR, LLR, and PESQ.

- The ICA–QPSO method produced better results than those of the ICA-GSO method according to the SNR, SIR and PESQ measurements.
- The proposed method utilised the following three objective functions: negentropy, differential entropy, and mutual information functions. The negentropy function generally produced the best results.
- The ICA–QPSO method with negentropy function demonstrated good accuracy results according to the SNR, SIR and PESQ measurements. Hence, this method is suitable for various applications that require accuracy.
- The ICA–PSO method with differential entropy and mutual information functions exhibited good results according to the measurements used in this study.

CONFLICT OF INTEREST

The authors have no conflict of relevant interest to this article.

REFERENCES

[1] Ondˇrej Sembera, Petr Tichavsky, and Zbynˇek Koldovsky, Adaptive Blind Separation of Instantaneous Linear Mixtures of Independent Sources, Springer International Publishing AG 2017, P. Tichavsk´ y et al. (Eds.): LVA/ICA 2017, LNCS 10169, pp. 172–181, 2017. DOI: 10.1007/978-3-319-53547-0 17.

[2] B. Paprocki, A. Pregowska, and J. Szczepanski, Optimizing information processing in brain-inspired neural networks, BULLETIN OF THE POLISH ACADEMY OF SCIENCES TECHNICAL SCIENCES, Vol. 68, No. 2, 2020.

[3] A. Hyvarinen, J. Karhunen, and E. Oja, "Independent Component Analysis", John Wiley & Son Inc., 2001.

[4] A. Cichocki and S. Amari, "Adaptive Blind Signal and Image processing learning algorithms and applications", John Willy & Sons, 2002.

[5] Ke Zhang, Yangjie Wei, Dan Wu and Yi Wang, Adaptive Speech Separation Based on Beamforming and Frequency Domain-Independent Component Analysis, Appl. Sci. 2020, 10, 2593; doi:10.3390/app10072593

[6] Ali Al-Saegh, Independent Component Analysis for Separation of Speech Mixtures: A Comparison Among Thirty Algorithms, Iraq J. Electrical and Electronic Engineering Vol.11 No.1 , 2015

[7] Y. Deville, "Blind Source Separation and Blind Mixture Identification Methods", J. Webster (ed.), Wiley Encyclopedia of Electrical and Electronics Engineering, John Wily & Sons Inc., 2016.

[8] Alaa Tharwat, Independent component analysis: An introduction, Applied Computing and Informatics, in press, 2018, https://doi.org/10.1016/j.aci.2018.08.006

[9] F. Rojas, C.G. Puntonet, I.Rojas, J. Ortega and A. Prieto, "Genetic Algorithm Approach to Nonlinear Blind Source Separation ", IEEE-2002.

[10] Liuyang Gao, Nae Zheng , Yinghua Tian, and Jingzhi Zhang, Target Signal Extraction Method Based on Enhanced ICA with Reference, Mathematical Problems in Engineering, Volume 2019, Article ID 4128438, https://doi.org/10.1155/2019/4128438.

[11] C.G. Puntonet, A. Mansourb, C. Bauerc, and E. Langc, "Separation of sources using simulated annealing and competitive learning", www.elsevier.com/locate/neucom, Neurocomputing 49 (2002) 39 – 60, 2002.

[12] J. Sun, C. Lai and X. Wu, "particle swarm optimization classical and quantum perspectives", CRC Press, 2012.

[13] B.K. Panigrahi, Y. Shi, and M.-H. Lim (Eds.), "Handbook of Swarm Intelligence", ALO 8, pp. 451–467.Springer-Verlag Berlin Heidelberg 2011.

[14] C. Liu, T. Sun, S. Hsieh, C. Lin, and K. Lee, "A Hybrid Blind Signal Separation Algorithm: Particle Swarm Optimization on Feed-Forward Neural Network", LNCS 4232, Springer-2006.

[15] K. Song, M. Ding, Q. Wang, and W. Liu, "Blind source separation in post-nonlinear mixtures using natural gradient descent and particle swarm optimization algorithm", LNCS 4493, Springer-2007.

[16] W. Yu, L. Zhenxing and L. ChangHai, "Improved particle swarm to nonlinear blind source separation", IEEE, 2007.

[17] T. Kurihara and K. Jin'no, "A Nonlinear blind source separation system using particle swarm optimization algorithm", Journal of Signal Processing, Japan, 2013.

[18] T. Lee and B. Koehler, "blind source separation of nonlinear mixing models", IEEE, 1997.

[19] A. Ziehe, M. Kawanabe, S. Harmeling and K. Muller, "Separation of post-nonlinear mixtures using ACE and temporal decorrelation", Proc. of the international conference on independent component analysis and signal separation (ICA 2001), USA, 2001.

[20] Z. Xiong and T. S. Huang, "Nonlinear independent component analysis (ICA) using power series and application to blind source separation", 2001.

[21] Y. Tan and J. Wang, "Nonlinear Blind Source Separation Using a Genetic Algorithm", 2001.

[22] J. Eriksson and V. Koivunen, "Blind identifiability of class of nonlinear instantaneous", Proc. of XI European signal processing conference (EUSIPCO 2002), France, 2002.

[23] J. Sole, C. Jutten and A. Taleb, "Parametric approach to blind deconvolution of nonlinear channels, Neuro-computing 48 (2002) 339–355, 2002.

[24] L. B. Almeida, "MISEP – Linear and Nonlinear ICA Based on Mutual Information", Journal of Machine Learning Research 4 (2003) 1297-1318, 2003.

[25] F. Rojas, C. G. Puntonet, M. Rodríguez-Álvarez, I. Rojas and R. Martín-Clemente, " Blind Source Separation In Post-Nonlinear Mixtures Using Competitive Learning, Simulated Annealing, And A Genetic Algorithm", IEEE Transactions On Systems, Man, And Cybernetics—Part C: Applications And Reviews, Vol. 34, No. 4, 2004.

[26] J. Karvonen and T. Tanaka, "Temporal Decorrelation as Preprocessing for Linear and Post-nonlinear ICA", Springer, C.G. Puntonet and A. Prieto (Eds.): ICA 2004, LNCS 3195, pp. 774–781, 2004.

[27] T. M. Dias, T. M. Dias, R. Attux, J. M.T. Romano and R.Suyama, "Blind Source Separation Of Post-Nonlinear Mixtures Using Evolutionary Computation And Gaussianization", Springer, 2009.

[28] F. Oveisi, S. Oveisi, A. Efranian, and I. Patras, "Nonlinear Independent Component Analysis for EEG-Based Brain-Computer Interface Systems", intechopen.com, 2012.

[29] B. Ehsandoust, M. Babaie-Zadeh, B. Rivet and C. Jutten, "Blind Source Separation in Nonlinear Mixtures: Separability and a Basic Algorithm", 1053-587X (c) 2016 IEEE, DOI 10.1109/TSP.2017.2708025, IEEE Transactions on Signal Processing, 2016.

[30] J. Kennedy and R. Eberhart," particle swarm optimization", Proceedings of the IEEE international conference on neural networks, Perth, Australia, pp. 1942-1948, 1995.

[31] J. Sun, B. Feng and W. Xu, "particle swarm optimization with particles having quantum behavior", IEEE-2004.

[32] W. Fang, J. Sun, Y. Ding, X. Wu and W. Xu, " A Review of Quantum-behaved Particle Swarm Optimization", IETE Technical Review, 27:4, 336-348, 2010.

[33] Z. Li and X. Huang, "Glowworm Swarm Optimization and Its Application to Blind Signal Separation", Hindawi Publishing Corporation, Mathematical Problems in Engineering, Volume 2016, Article ID 5481602.

[34] Philip Coleman, Qingju Liu, Jon Francombe, and Philip J. B. Jackson, Perceptual Evaluation of Blind Source Separation in Object-Based Audio Production, Springer International Publishing AG, part of Springer Nature 2018, Y. Deville et al. (Eds.): LVA/ICA 2018, LNCS 10891, pp. 558–567, 2018. https://doi.org/10.1007/978-3-319-93764-9_51

[35] Nidaa A. Abbas, Hussein Mohammed Salman, “Independent component analysis based on quantum particle swarm optimization”, Egyptian Informatics Journal, Vol. 19, Issue 2, July 2018, pp. 101-105, Publisher Elsevier.

[36] Nidaa A. Abbas, Jahanshah Kabudian, "Speech Scrambling based on Independent Component Analysis and Particle Swarm Optimization", International Arab Journal of Information Technology (IAJIT), impact factor 0.519, Vol.14, No.4, 2017.

[37] http://www.utdallas.edu/~loizou/speech/noizeus/