Post-processing of Time-Frequency Representations in Instantaneous Frequency Estimation Based on Ant Colony Optimization

Miloš Brajović, Vesna Popović-Bugarin*, Igor Djurović, Slobodan Djukanović

University of Montenegro/Faculty of Electrical Engineering
20 000, Podgorica, Montenegro
phone: + (382) 20 245 839, fax: + (382) 20 245 873

Abstract

Instantaneous frequency (IF) estimation of non-stationary signals embedded in high noise is addressed. When present, high noise represents a dominant error source in the IF estimation. Additive Gaussian noise with variance proportional to the signal power is assumed. An estimation approach based on the ant colony optimization (ACO) and time-frequency (TF) analysis is proposed. The ACO algorithm is adapted for the IF estimation starting from the Wigner distribution (WD) of the considered signal. The proposed technique is also applicable to numerous other representations, without any change in the parameter setup. This method surpasses the influence of high noise in the IF estimation; for that purpose, we have designed a pheromone deposition gradient and a mechanism for the variation of the agents' population size. The introduced approach improves the fast-varying IF estimation accuracy, overcoming known issues in the state-of-the-art algorithms dealing with high noise. The basic principles of the proposed method are illustrated and performance validated through numerical examples.

Keywords: Ant colony optimization, instantaneous frequency estimation, time-frequency signal analysis, Wigner distribution

1. Introduction

Time-frequency (TF) signal analysis has drawn a significant attention during the last few decades [1]-[43]. Various TF representations enable efficient analysis and extraction of information contained within time variations of the signal's spectral content [4], [6]. One of the key topics in this area is the instantaneous frequency (IF) estimation of a signal [4]-[10]. Numerous TF representations and algorithms have been proposed in order to facilitate the IF estimation, but no TF representation exists, nor the estimation algorithm, that resolves this problem for all classes of

* Corresponding author
Email address: pvesna@ac.me (Vesna Popović-Bugarin)

signals and under all circumstances [4], [5]. These facts make the IF estimation problem still scientifically attractive [6].

The IF estimation arises in numerous application fields including communications based on frequency modulation (FM), radar and sonar systems, speech analysis and recognition, analysis of video signals, seismology, biology, bio-medicine [6], [16], [25].

Many TF representations have the property of concentrating the signal energy at and around the IF [15], [29]-[31]. This is the reason why the IF estimation formulation, in classical estimation approaches, reduces to determination of the TF representation maxima [4], [5], [6]. The Wigner distribution (WD) has been widely used as an IF estimator of FM signals, since the signal representation in the TF plane is highly concentrated [8], [26], [27]. However, due to higher-order signal phase derivatives, the WD contains inner interferences. Moreover, when multicomponent signals are considered, the undesired cross-terms appear in addition to signal components referred to as the auto-terms. The cross-terms can mask the auto-terms in the TF plane [3], [25]. Besides the principles presented in [4] and [6], a comprehensive analysis of the WD as an IF estimator is given in [7] and [8], where the estimation error sources were classified into four categories: bias, errors due to variations within the signal's auto-terms, errors due to frequency discretization and errors due to high noise. Techniques that deal with the first three error sources are presented in [4], [9], [27], [28]. The influence of high noise has attracted a significant attention, since it represents a dominant error source when it occurs [7]-[9]. Errors due to high noise appear since the high noise induces false maxima (maxima outside of the auto-terms) in the TF plane. Under the term "high noise" we assume additive Gaussian noise of the constant variance proportional to the power of the contaminated signal, as considered in [6]-[9], [14].

An instance of the Viterbi algorithm (VA), originally introduced in [8], has been applied in the IF estimation in order to overcome the negative high noise impact. The performance of the VA-based approach in various estimation problems involving high noise has been confirmed during the recent years [6], [9]-[13]. Another IF estimation approach based on a high-dimensional search of the IF curves in the TF plane has been proposed in [30], for the case of wavelets. The search of curves is based on a stochastic relaxation procedure. This approach has certain robustness to high noise. The generalization of this approach towards the TF representations with an introduction of a new stochastic search procedure has been presented in [31].

Artificial ant colonies (AACs), a biologically inspired paradigm, represent multi-agent tools for problem solving without a centralized control [44]-[59]. The whole set of optimization techniques based on the AAC concept, widely known as the ant colony optimization (ACO), has been developed and applied in different scientific areas, especially

where the hard-solving local optimization problems arise [46]-[56]. The AACs are one of many concepts in the so-called swarm intelligence, where a population of artificial agents forms a collective intelligence over a specific environment [44]. Important application fields in digital image processing include edge detection, pattern recognition and feature extraction [48]-[56]. For the problem considered in this paper, an especially interesting ACO application is in edge detection in digital images [49], [54], [56]. By taking into account that edges represent image segments with high contrast and/or color variations, the median difference in agents' neighboring points has been used as the main criterion for edge detection [50], [54], [56].

In this paper, the ACO algorithms proposed in [49] and [50] are adapted for the TF-based IF estimation. A new gradient which takes into account the IF properties is introduced in order to achieve a robust estimation in high noise: the IF should pass through as many as possible points of the TFR with highest magnitudes, while the IF variation between two consecutive points should not be too fast [8].

The estimation is performed based on a generated pheromone map, representing a new TF representation with significantly reduced disturbances. The initial version of the algorithm is proposed in [59]. In this paper, we improve the performance of the algorithm [59] by introducing variable population size [49], [50]. In addition, we evaluate the performance of the proposed method versus signal-to-noise ratio (SNR), provide comparison with the state-of-the-art VA-based algorithm and illustration of cross-terms suppression in the case of multicomponent signals, as well as the IF estimation illustration for real-life signals. The concept of variable population size provides an additional control mechanism for the mass behavior of an AAC. The basic idea is to retain the agents moving across the TF points corresponding to the auto-terms, while tending to eliminate as many those not corresponding to the auto-terms as possible. The agent's ability to survive during the iterations is measured by its energy, which is related to the proposed gradient. In this way, the influence of the gradient on the mass behavior of agents is emphasized. The distributed nature of the proposed algorithm and carefully designed gradient make the estimation more robust to fast IF variations. In this way, it overcomes, at a certain level, the sensitivity of the VA to fast IF variations, which is confirmed for signals with fast IF variations. Additionally, the proposed approach suppresses inner interferences, as well as cross-terms in multi-component signals.

Basic theory concerning the IF estimation problem is given in Section 2. The ACO algorithm within the framework of the TF-based IF estimation is presented in Section 3. Section 3 also introduces the pheromone deposition gradient and the variable population concept, both adapted for the IF estimation, the estimation algorithm and a discussion on the algorithm parameters. Section 4 presents numerical examples with the estimation results and a statistical

verification of the proposed method by a comparison with the WD maxima and the VA estimators. Section 5 concludes the paper.

2. Background theory

Consider a complex amplitude and frequency modulated signal [6]-[9]:

$$s(t) = A(t)e^{j\phi(t)}, \tag{1}$$

where $A(t)$ is a slowly varying amplitude with respect to phase variations, i.e. $|dA(t)/dt| \ll |d\phi(t)/dt|$ and $\phi(t)$ is the signal instantaneous phase. Note that for real-valued signals positive frequencies are taken into account or the corresponding Hilbert transform of the signal is calculated. Some alternative signal models are also discussed in the literature [17]-[22]. The IF of $s(t)$ is defined as the first derivative of its phase, i.e.

$$\Omega(t) = d\phi(t)/dt. \tag{2}$$

The signal of interest is embedded in complex additive white Gaussian noise (AWGN) $\varepsilon(t)$ with zero-mean and variance σ^2, with independent and identically distributed (i.i.d.) real and imaginary parts, that is

$$x(t) = s(t) + \varepsilon(t). \tag{3}$$

The signal is sampled with the sampling interval Δt to obtain $x(n) = x(n\Delta t)$, where n represents discrete time variable. In the case when $s(t)$ contains multiple components, i.e. when $x(t)$ can be written as

$$x(t) = \sum_{l=1}^{S} s_l(t) + \varepsilon(t), \tag{4}$$

where S represents the number of components $s_l(t)$ defined by (1), the IF of each component can be calculated as the first derivative of the corresponding phase component. This concept has not full theoretical foundation, but it has clear justification when components are well-separated in the TF plane.

Further, we observe the WD of the signal $x(n)$, well known for its advantageous properties in the IF estimation of noisy signals. The IF approach that we present may be applied on other TF representations as well.

The WD of $x(n)$ [8] is defined as

$$WD(n,k) = \sum_{m=-K/2}^{K/2-1} w(m)x(n+m)x^*(n-m)e^{-j4\pi mk/K}, \tag{5}$$

where $-K/2 \le k \le K/2-1$ represents discrete frequency index, K denotes the length of a real-valued symmetric window $w(n)$, whereas $*$ denotes complex conjugation. Without loss of generality even K is assumed. It is also assumed that the

discrete signal length is N_s. For a given instant n, the IF ($\omega(n)$ or $\Omega(n\Delta t) = \omega(n)/\Delta t$) is estimated as the WD maximum position [6], [8], [25]:

$$\hat{k} = \arg\max_k WD(n,k). \tag{6}$$

This simple IF estimator is a common tool in practice. However, a high noise causes the WD maxima to be located away from the IF points, thus resulting in erroneous IF estimation [7], [8].

The state-of-the-art approach overcoming the IF estimation problem in high noise is an instance of VA, originally introduced in [8], whose performance is reviewed in [6]. This algorithm combines a non-parametric method based on the WD maxima with the minimization of IF variations between consecutive points. As the VA incorporates a criterion that assumes slow IF variations between consecutive points, it is sensitive to fast IF variations.

3. IF estimation by using Ant Colony Optimization algorithm

3.1 The review of basic ACO concepts, notation and terminology

Basic ACO algorithm is explained in [46], while details on ACO applications in edge detection and feature extraction can be found in [49]-[53]. The discrete WD, i.e. a matrix of size $N \times K$, where $N = N_s - K$ can be observed as a rectangular grid in which artificial ants move across adjacent cells. In other words, we observe the domain consisted of TF points $(n,k) \in [0,N) \times [-K/2, K/2)$. Ants are initially placed at random grid positions, with random orientations.

Transition of an ant from cell $(n',k') \in [0,N) \times [-K/2, K/2)$ to a new cell (n, k) occurs in each iteration I. An iteration is finished when each ant from the colony moves from a previous position (n',k') to allowed neighboring position (n, k). In every cell visited by an ant a certain pheromone level is added. Pheromone level deposited at the position (n, k) is denoted by $\Phi(n, k)$. A matrix known as the pheromone map $\mathbf{\Phi}$ is used to store the pheromone level for every cell (position) in the observed grid. In this paper, the pheromone map represents a new TF representation, which is used for the IF estimation. When a predefined number of iterations is reached, or another stopping criterion is met, the algorithm terminates.

The general ant transition rule from the current cell (n',k') to the new cell (n, k) in iteration I, is defined by probability $P^{(I)}_{(n',k')}(n,k)$, which depends on the pheromone level $\Phi(n, k)$, and a heuristic function $d_{(n',k')}(n,k)$, [46], [50]-[53]. Herein, we use the adaptation of the basic ACO algorithm [46] for the case of digital images, presented in [50], where this probability is defined as:

$$P_{(n',k')}^{(I)}(n,k)=\begin{cases}\dfrac{\left(\Phi^{(I)}(n,k)\right)^{\beta} d_{(n',k')}(n,k)^{\zeta}}{\sum_{(n,k)\in\mathbf{Q}(n',k')}\left(\Phi^{(I)}(n,k)\right)^{\beta} d_{(n',k')}(n,k)^{\zeta}}, & (n,k)\in\mathbf{Q}(n',k') \\ 0, & \text{otherwise,}\end{cases} \tag{7}$$

where $\mathbf{Q}(n',k')$ is the set representing the permissible ant's movement range, i.e. the local neighborhood of the current ant's position (n',k'); parameters β and ζ control the influence of pheromone and heuristic information, respectively [46], [50], [52]. As it is emphasized in [45], instead of $\left(\Phi^{(I)}(n,k)\right)^{\beta}$ and $d_{(n',k')}(n,k)^{\zeta}$ other monotonic non-decreasing functions $g(\Phi(n,k))$ of the pheromone level can be used, and several of them are proposed in the literature, depending on applications [45]-[55]. In this paper, we fix $\zeta=1$ and adapt the original ACO algorithm via a pheromone update rule. In the end of each iteration, the pheromone decreases globally in the whole pheromone map, thus allowing the control of the pheromone map content by a suitably defined pheromone update rule.

3.2 Ant Colony Optimization

3.2.1 Initialization

At the beginning of the optimization algorithm, γ percents of the TF representation points are randomly covered with intelligent agents (ants). The initial ant positions $(n_0,k_0)\in[0,N)\times[-K/2,K/2)$ are obtained by a random number generator with uniform distribution. The initial ant orientations are defined by values $P(n, k)$ stored in an auxiliary matrix **P** with elements:

$$P(n,k)=\begin{cases}r, & (n,k)=(n_0,k_0)\\ 0, & (n,k)\neq(n_0,k_0)\end{cases} \tag{8}$$

where r is a random integer with uniform distribution in the range {1, 2,…, 8}. Each r value corresponds to one of eight possible orientations (1 denotes the upwards orientation, 2 left upwards, 3 left etc.). In all subsequent iterations, the ant orientation for the new cell (n, k) is determined by its movement from the old cell (n',k'). For example, if an ant moves from $(n',k')=(n+1, k)$ to (n, k), then it is upwards oriented, i.e., $P(n, k)$ equals 1. When ant leaves (n',k'), the corresponding value in the matrix **P** is set to zero.

Ants communicate with each other via a crucial concept of pheromone deposition and evaporation. Initially, the elements of the pheromone map at the positions occupied by ants are set to a small positive constant $\Phi_0 < 1$:

$$\Phi^{(0)}(n,k)=\begin{cases}\Phi_0, & (n,k)=(n_0,k_0)\\ 0, & (n,k)\neq(n_0,k_0).\end{cases} \tag{9}$$

The ant movement and pheromone deposition mechanisms are crucial for the control of mass behavior of artificial ants [45]-[56].

3.2.2 Ant transition rule

After the initialization, in each algorithm iteration ants move according to certain rules. Let us observe one ant's transition from cell (n',k') to a neighboring cell $(n,k)\in \mathbf{Q}(n',k')$. Every agent can move only to adjacent cells (positions), depositing a certain level of pheromone on that cell. One cell can be occupied by only one ant, and ants do not move if they are completely surrounded by other ants. An iteration ends when all the ants move to adjacent cells (except the totally surrounded ones).

An ant chooses a cell to move to based on two criteria: the pheromone level in adjacent cells [46], [50] and its current orientation [53]-[55]. Since a discrete rectangular grid is considered, an ant located at the position (n',k') in the *I*-th iteration can move to one of eight adjacent cells (*n*, *k*) belonging to the set $\mathbf{Q}(n',k')$. An ant at the position (n',k') is shown in Fig. 1, where the corresponding neighboring cells $(n,k)\in \mathbf{Q}(n',k')$ are shaded.

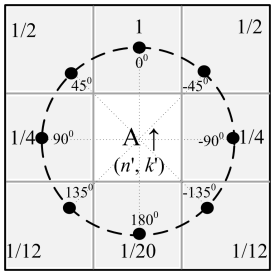

Fig. 1 Agent oriented upwards with $P(n',k')=1$ and corresponding values of $d\left(\Theta_{(n',k')}(n,k)\right)$ for possible discrete directions – angles (dots). The permissible ant's movement range $\mathbf{Q}(n',k')$ is shaded.

An ant at the original position (n',k') has a certain orientation $P(n',k')$ according to (8), that is, the orientation of the point (n',k') towards the adjacent cells $(n,k)\in \mathbf{Q}(n',k')$ as depicted in Fig. 1. Values of $P(n',k')$ are fixed, $P(n',k')\in\{1,2,...,8\}$, where $P(n',k')=1$ denotes an upward oriented ant, while other numbers correspond to the orientations towards the remaining 7 cells in counterclockwise direction. The orientation has an important role in the choice of the targeting cell $(n,k)\in \mathbf{Q}(n',k')$. The angles between an upwards oriented ant at (n',k') and the adjacent cells positions are illustrated in Fig. 1. As in [54], a discrete variable $\Theta_1(n,k)\in\{0^o,\pm45^o,\pm90^o,\pm135^o,180^o\}$ is introduced, depending only on the position of the cell $(n,k)\in \mathbf{Q}(n',k')$ relative to the cell (n',k') corresponding to the

ant's orientation shown in Fig. 1. The corresponding angles $\Theta_{(n',k')}(n,k)$ for other possible ant's orientations can be calculated as:

$$\Theta_{(n',k')}(n,k) = \Theta_1(n,k) - \left[P(n',k') - 1\right] \cdot 45^{\circ} \tag{10}$$

where $\Theta_{(n',k')}(n,k)$ represents the angle between the current ant's orientation and the adjacent cell.

Depending on the application, different heuristic functions $d_{(n',k')}(n,k)$ used for defining the probability (7) have been proposed [45]-[59]. In this paper, we use the heuristic function $d\left(\Theta_{(n',k')}(n,k)\right)$ as defined in [49], [53]-[55], i.e. the function of angles $\Theta_{(n',k')}(n,k)$ calculated as:

$$d\left(\Theta_{(n',k')}(n,k)\right) = \begin{cases} 1, & \Theta_{(n',k')}(n,k) = 0^{\circ} \\ 1/2, & \Theta_{(n',k')}(n,k) = \pm 45^{\circ} \\ 1/4, & \Theta_{(n',k')}(n,k) = \pm 90^{\circ} \\ 1/12, & \Theta_{(n',k')}(n,k) = \pm 135^{\circ} \\ 1/20, & \Theta_{(n',k')}(n,k) = 180^{\circ}, \end{cases} \tag{11}$$

for $(n,k) \in \mathbf{Q}(n',k')$.

Heuristic function $d\left(\Theta_{(n',k')}(n,k)\right)$ in (11) is defined so that agents most likely preserve the orientation from the previous iteration, thus $d\left(\Theta_{(n',k')}(n,k)\right)$ has the largest value for angle $\Theta(n,k) = 0^{\circ}$. With the increase of the angle, the value of $d\left(\Theta_{(n',k')}(n,k)\right)$ gradually decreases, also decreasing the probability that ant moves to these directions, as it can be seen in (7). In this way, very sharp turns are much less likely than turns through smaller angles [53]-[55], which is also in accordance with the IF estimation problem [6]-[9], [25].

The second parameter that influences the movement of an ant, and which can be used for the ant colony behavior control, is dependence on the pheromone level $\Phi(n,k)$, given by a function defined as [46]-[54]:

$$g\left(\Phi(n,k)\right) = \left(1 + \frac{\Phi(n,k)}{1 + \delta\Phi(n,k)}\right)^{\beta}. \tag{12}$$

A large value of β results in ants heavily attracted by the pheromone level and vice versa. The parameter δ controls the ant's sensitivity to the pheromone concentration [53]-[56]. With the introduced function $g\left(\Phi(n,k)\right)$, the probability (7) is now given by [46]-[49]:

$$P_{(n',k')}^{(I)}(n,k)=\begin{cases}\dfrac{g\left(\Phi^{(I)}(n,k)\right)d\left(\Theta_{(n',k')}(n,k)\right)}{\sum_{(n,k)\in\mathbf{Q}(n',k')}g\left(\Phi^{(I)}(n,k)\right)d\left(\Theta_{(n',k')}(n,k)\right)}, & (n,k)\in\mathbf{Q}(n',k')\\ 0, & \text{otherwise.}\end{cases} \tag{13}$$

3.2.3 Pheromone update

When an ant moves to a neighboring cell $(n,k)\in\mathbf{Q}(n',k')$, the pheromone level is updated according to [46]-[49]:

$$\Phi^{(I+1)}(n,k)=\Phi^{(I)}(n,k)+\mu\nabla(n,k)+\xi, \tag{14}$$

with ξ being a small constant level of pheromone, $\nabla(n,k)$ is the gradient, i.e. a dynamic pheromone value which is added to new position (n, k) visited by the agent and μ is a positive step that controls the pheromone amount added by $\nabla(n,k)$. The fixed pheromone amount ξ is added in order to ensure that, after the pheromone evaporation at the end of the current iteration I by value ξ, all visited positions have the pheromone level at least equal to the gradient $\nabla(n,k)$ value [53]-[55]. For the IF estimation, the gradient $\nabla(n,k)$ is defined taking into account the specific nature of the considered problem. It is analyzed in the following subsection.

At the end of each iteration a constant pheromone amount ξ evaporates from all cells, as it is discussed in detail in [45]-[56], i.e. for each (n, k) the pheromone map is updated according to

$$\Phi^{(I+1)}(n,k)=\begin{cases}\Phi^{(I)}(n,k)-\xi, & \Phi^{(I)}(n,k)\geq\xi\\ 0, & \Phi^{(I)}(n,k)<\xi.\end{cases} \tag{15}$$

If the new pheromone level is negative, then it is set to zero.

The previously described algorithm is summarized in Appendix A.

3.3 Pheromone deposition gradient in the IF estimation

For the definition of gradient, the nature of the considered problem, i.e., the IF estimation, has to be taken into account. The IF estimate should be extracted from the pheromone map, obtained after the algorithm has finished. To this aim, we use the fact that although the WD maxima in a high noise environment are likely to be dislocated from the true IF of the signal [7], at each time instant n one of the largest WD values will still be positioned at the IF [7], [9], [25]. Dislocation is a consequence of high-level noise peaks which surpass the WD values inside the auto-terms. On the other side, it is known that IF variation between two consecutive time instants should not be large, which is the most common case in real applications [8]. Taking into account these two facts, we introduce a new gradient $\nabla(n,k)$ form, which defines the pheromone gradient in (14)

$$\nabla(n,k) = \Psi(n,k)\Xi(n,k)\Lambda(n,k). \tag{16}$$

We define the function $\Psi(n, k)$ as

$$\Psi(n,k) = \frac{1}{27}\prod_{i=-1}^{1}\sum_{j=-1}^{1} WD(n+j,k+i), \tag{17}$$

i.e. as the product of the mean values of columns of the 3×3 adjacency in the WD matrix, centered at point (n, k). Namely, if an auto-term appears within the observed 3×3 adjacency of the point (n, k) then a large value of $\Psi(n, k)$ is expected. It is expected that an auto-term appears in all three columns of the 3×3 adjacency, as shown in Fig 2. (*a*)-(*f*). On the other side, if a 3×3 adjacency contains only noise, then $\Psi(n, k)$ has a small value, since high-level noisy WD points are usually isolated in the TF plane. Hence, it is expected that only few high-level noisy points exist in the local neighborhood of (n, k), whereas other points are of lower level (Fig. 3 (*a*)-(*f*)). The function $\Psi(n, k)$, whose values are presented in Tables 1 and 2, has high values for the matrices from Fig. 2, containing the WD auto-terms, and lower values for the matrices in Fig. 3, which represents 3×3 neighborhoods of point (n, k) placed outside the auto-terms, containing noise only. In this way, $\Psi(n, k)$ emphasizes the auto-terms presence in the gradient calculation (16).

Table 1

Values of functions (17), (18) and (19) for 3 x 3 matrices shown in Fig. 2.

	(*a*)	(*b*)	(*c*)	(*d*)	(*e*)	(*f*)
$\Psi(n,k)$	60310	21021	25842	17167	48840	8911
$\Xi(n,k)$	40.44	27.67	30.22	27.56	36.67	20.78
$\Lambda(n,k)$	151032	114070	171000	193662	209560	204435

Table 2

Values of functions (17), (18) and (19) for 3 x 3 matrices shown in Fig. 3.

	(*a*)	(*b*)	(*c*)	(*d*)	(*e*)	(*f*)
$\Psi(n,k)$	4344.9	-29.63	-10920	0	63.33	3422.2
$\Xi(n,k)$	7.11	-0.78	5	13.44	-11.33	23
$\Lambda(n,k)$	10488	660	29140	2440	17496	12400

We define $\Xi(n, k)$ as the mean value of the observed 3×3 adjacency, i.e.

$$\Xi(n,k) = \frac{1}{9}\sum_{i=-1}^{1}\sum_{j=-1}^{1} WD(n+j,k+i). \tag{18}$$

It is expected that $\Xi(n, k)$ has small values for noisy WD points (outside of the auto-terms), since high-level noise peaks are usually isolated [6]-[9]. On the other hand, due to the windowing effects in the WD calculation and the fact that only linear FM (LFM) signals are ideally concentrated in the WD, it is expected that if the local 3×3 adjacency centered at point (n, k) contains an auto-term, it will occupy three or more points in the neighborhood. If the signal is not ideally concentrated, even the whole observed 3×3 WD sub-matrix may be occupied by an auto-term, thus producing a large $\Xi(n, k)$ value. Therefore, the largest values of $\Xi(n, k)$ are expected to occur for auto-terms. Note that the WD values are Gaussian in nature, with non-zero mean within the auto-terms and zero mean outside the auto-terms [7].

Now consider again the matrices shown in Figs. 2 and 3. The corresponding values of $\Xi(n, k)$ are shown in Tables 1 and 2 (second rows). This function mainly produces higher values for the matrices from Fig. 2 than those from Fig. 3. The exception is the matrix shown in Fig. 3 (*f*), whose $\Xi(n, k)$ value exceeds that of the matrix shown in Fig. 2 (*f*). This problem is solved by incorporating the function $\Lambda(n, k)$ in (16), which is shown to output significantly higher values for the matrix in Fig. 2 (*f*) than for that in Fig. 3 (*f*). The function $\Lambda(n, k)$ takes into account the fact that the IF has small variations at consecutive time instants *n,* also in the 3×3 neighborhood of the point $WD(n, k)$:

$$\Lambda(n,k)=\max\left(\left[\prod_{i=-1}^{1}WD(n+i,k+i)\quad\prod_{i=-1}^{1}WD(n+i,k-i)\quad\prod_{i=-1}^{1}WD(n+i,k-1)\quad\prod_{i=-1}^{1}WD(n+i,k)\quad\prod_{i=-1}^{1}WD(n+i,k+1)\right]\right). \tag{19}$$

The first term is the product of the elements on the main diagonal of a 3×3 neighborhood of the point $WD(n, k)$, whereas the second term is the element product on the secondary diagonal of this neighborhood; third, fourth and fifth elements are element products in the first, second and third rows of the neighborhood, respectively. The heuristic function (19) takes the maximal value of the vector consisted of these products.

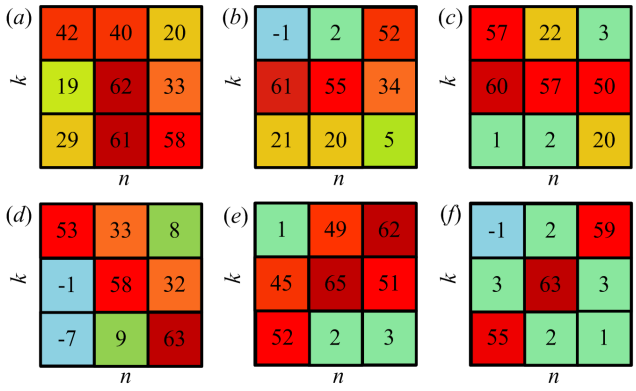

Fig. 2 Illustration of the 3×3 neighborhoods of (n, k), which correspond to auto-term positions. The WD values have been rounded to the nearest integer for illustration clarity.

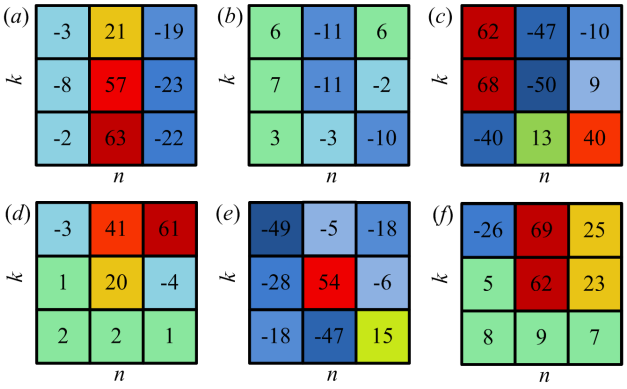

Fig. 3 Illustration of the 3×3 neighborhoods of (n, k), which correspond to non-IF positions (noise). The WD values have been rounded to the nearest integer for illustration clarity.

If the WD values corresponding to the IF are within the observed 3×3 neighborhood of the point $WD(n, k)$, at least one of the products in (19) is expected to have a large value. On the other side, one or two isolated WD points with a large value corresponding to the noise will probably produce smaller $\Lambda(n, k)$ values.

Table 3

Gradient values (16) calculated for matrices shown in Fig. 2 (second row) and Fig. 3 (third row). Values are scaled by 10^{10} for better presentation.

	(*a*)	(*b*)	(*c*)	(*d*)	(*e*)	(*f*)
$\nabla_1(n,k)/10^{10}$	36.84	6.63	13.36	9.16	37.53	3.78
$\nabla_2(n,k)/10^{10}$	0.0324	0	-0.1591	0	-0.0013	0.0976

The values of $\Lambda(n, k)$, calculated for 3×3 neighborhoods of the point (n, k) given in Fig. 2 and Fig. 3, are shown in Tables 1 and 2 (third rows). Clearly, this function gives preferences to matrices with large values distributed along rows or diagonals of the matrix. These shapes most likely correspond to the IF points (auto-terms).

Finally, the heuristic functions product in (16) results in a large value for the neighborhood of auto-terms. Table 3 summarizes the gradient values for the matrices in Fig. 2, denoted as $\nabla_1(n,k)$, and the matrices in Fig. 3, denoted as $\nabla_2(n,k)$. As expected, $\nabla_1(n,k)$ exceeds $\nabla_2(n,k)$ to a great extent in all considered cases.

The summary of the proposed algorithm for the pheromone map generation is given in Appendix B.

3.4 Population size variation

In the literature, methods for the variation of agents population based on both aging/dying and reproduction process, i.e. positive and negative feedback are presented [49], [53]-[55]. Our adaptation consists of the elimination of the agents which move through undesirable positions, i.e. points where the pheromone level is sufficiently low during

a certain number of iterations. In this way, ants which are too far from the TF points corresponding to the auto-terms are eliminated.

The IF estimation is performed using the pheromone map **Φ**. Since the pheromone deposition control is established through the gradient (16) in (14) for each ant independently and locally, our aim is to find a method for additional control of the ants number, based on the pheromone amount which they have left during the previous iterations. In the considered framework, it is important to eliminate as many undesirable agents as possible, since their pheromone trace usually causes large estimation errors. Namely, the idea is to give an advantage to ants that often move to cells corresponding to the auto-terms, depositing more pheromone in these cells. This can be measured with the gradient $\nabla(n,k)$. The population reduction concepts are implemented as presented below.

At the beginning of the algorithm described in Appendix A, in Step 0, a fixed level of energy $1+\alpha$ is assigned to every ant, stored in auxiliary matrix **E**:

$$E^{(0)}(n,k)=\begin{cases}1+\alpha, & (n,k)=(n_0,k_0)\\ 0, & (n,k)\neq(n_0,k_0).\end{cases} \tag{20}$$

The parameter $\alpha < 1$ is a small positive constant. The initial α value gives agents a chance to survive during at least two following iterations [49], [53]-[55].

Let us observe again one ant's transition from cell (n',k') to a neighboring cell $(n,k)\in\mathbf{Q}(n',k')$. In Appendix A, Step 2, the energy corresponding to new position determined by the probability (13) is updated according to [49]-[54]:

$$E^{(I+1)}(n,k)=E^{(I)}(n',k')+\alpha\mu\nabla(n,k), \tag{21}$$

where the gradient $\nabla(n,k)$ calculated in (16) is used. Then, the energy matrix **E** is updated by setting the element corresponding to the previous ant position (n', k') to zero, i.e.:

$$E^{(I+1)}(n',k')=0. \tag{22}$$

In Step 3, the energy level of all ants kept in the matrix **E** is decreased by α, for every $(n,k)\in[0,N)\times[-K/2,K/2)$

$$E^{(I+1)}(n,k)=E^{(I)}(n,k)-\alpha. \tag{23}$$

In this step, all ants at positions with energy $E(n,k)\leq 0$ are eliminated by removing the corresponding elements of the matrix **P**, for each $(n,k)\in[0,N)\times[-K/2,K/2)$:

$$P^{(I+1)}(n,k)=0 \text{ if } E^{(I+1)}(n,k)<0. \tag{24}$$

Since the population of agents is decreased during the iterations when the IF is detectable, the stopping criterion of the algorithm can be the minimal permitted number of agents, obtained from the matrix **P** as the number of non-zero elements. In the case of ideally concentrated TF representation of a non-noisy mono-component signal, the IF will have exactly N non-zero values, i.e. one value for each time instant n, if it is defined for every N. However, if we take into account that in a high noise environment the algorithm can remove a certain number of auto-term points, as well as the fact that, in general, signals are not ideally concentrated, the algorithm can be stopped when the number of agents falls below 80-100% of N.

3.5 Parameter discussion

Determination of the optimal ACO algorithm setup is a complex task [44]-[59], especially in applications such as image processing, due to a large number of parameters [49]-[56]. The same holds for the application of ACO in the IF estimation problem. It is also important to emphasize that the parameter setting of ACO could not be mathematically derived [50]. Hence, in this paper, as it is usual in ACO applications [44]-[59], a part of suitable parameters is determined empirically, whereas the certain number of them is set based on detailed analysis and numerical results in the literature [49]-[56], as it is presented in the sequel. In Table 4, the set of parameters used to obtain numerical results in the paper is given. In Section 4.4 some numerical results regarding the parameter selection are presented.

Table 4

Parameter values used in the proposed approach. The Range row represents the ranges of tested values of some parameters that produced satisfactory results in numerical examples.

Parameter	μ	γ	ξ	α	Φ_0	β	δ
Value	$2/L^4$	70%	0.07	0.7	0.2	3.5	0.2
Range	$\mu \geq 1/L^4$	$50\% \leq \gamma \leq 90\%$	$0 < \xi < 0.1$	$0 < \alpha < 1$	$0.1 \leq \Phi_0 \leq 0.2$		

1. Step μ represents one of the most important parameters in the pheromone update (14) and energy update (21). In the proposed algorithm, the μ value is heuristically determined. It is in fact a normalization factor for the proposed gradient (16), and thus it has to be set in accordance with the gradient structure. If the step is too small, the influence of $\nabla(n,k)$ is not strong enough to dictate the pheromone deposition on the auto-term points.

   Let us denote with c the maximal value of WD(n, k) 3 x 3 neighborhood, when it belongs to an auto-term. If we analyze the functions (17), (18) and (19), it follows that $\Psi(n, k) \sim c^3$, $\Xi(n, k) \sim c$ and $\Lambda(n,k) \sim c^3$. Therefore, when the

gradient "detects" an auto-term, it is proportional to c^7. Generally, c is not constant due to amplitude variations and noise influence. In addition, it is generally smaller than the global TF maximum

$$L = \max_{(n,k)} \{WD(n,k)\}, (n,k) \in [0,N) \times [-K/2, K/2). \tag{25}$$

Hence, we propose to choose the step μ in (14) and (21), used for scaling the gradient $\nabla(n,k)$, as follows

$$\mu \geq \frac{1}{L^4}. \tag{26}$$

Note that although the same fixed value of the step $\mu = 2/L^4$ is used for all presented numerical results, a more suitable choice of this parameter may lead to better estimation results of the proposed approach. The condition (26) is experimentally evaluated as shown in Section 4.4.

2. The next important parameter is the initial population size, defined as the percentage γ of the product $N \times K$, where N and K are the TF grid dimensions. If the initial population is set too small, there is a risk that agents will be eliminated from the TF plane before they even approach the auto-terms. Moreover, there is a risk that, due to the auto-terms corruption in a high noise environment, only a small number of ants will reach the auto-terms, deposit pheromone according to (14) and consequently attract other ants. In addition, the population control using (21) would not be efficient. In this case, it is reasonable to choose a population size which is not too low. We have analyzed the performance of the proposed method with various values of γ and it has been shown (Section 4.4) that $0.5 \leq \gamma \leq 0.9$ cannot endanger the algorithm output. Thus, in this paper $\gamma = 0.7$ is used for all presented results.
3. The constant ξ is known as the pheromone evaporation rate, which is used to avoid limitless deposition of the pheromone trails and restrain the ants from choosing the same cells [50]. It is also the part of the pheromone update rule (14), ensuring that all positions visited by ants have the pheromone level at least proportional to the normalized gradient after the evaporation. Our numerical results indicate that $\xi < 0.1$ should be used.

   The pheromone constant Φ_0 used for the initialization of the pheromone matrix in (9) can be set such that $\Phi_0 \leq 0.2$. In this way, during the first few iterations pheromone level still remains positive in these cells after the evaporation. We have tested many proposed values of this parameter [49]-[54] with different signals and different SNRs, and our experimental results have shown that in the range of values presented in the literature $0.1 \leq \Phi_0 \leq 2$ algorithm is quite robust to the selection of this parameter at lower SNRs, whereas it should have smaller values $\Phi_0 \leq 0.2$ at higher SNR values.

We have found suitable to use values $\xi = 0.07$ and $\Phi_0 = 0.2$ obtained empirically and confirmed numerically as presented in the literature [49], [53], [54].

4. Parameter α is used for the energy update (21) and for the energy decrease in (23), similarly as the parameter ξ is used in the pheromone update process. Parameter α is originally introduced in [49], [53], [54]. However, unlike the pheromone that is related to a position in the pheromone map, the energy is associated with the observed ant. According to (21), when an ant moves, the energy matrix is updated in a way that the energy in new agent's position (n, k) is calculated based on the energy in previous agent's position (n', k'). This fact explains the need for new parameter, different from ξ. This parameter dictates the rate of energy decrease, and the dynamics of ant dying process. A too large value leads to the elimination of all agents. However, note that in order to prevent this, we have proposed that the algorithm terminates not only when the maximal allowed number of iterations $I_{\max}$ is reached, but also when the number of agents falls below 100% of N, to ensure that a reasonable number of agents remain to perform the estimation. Since the gradient normalized with the parameter μ is used in (21), we experimentally determined value of the parameter as $\alpha = 0.7$. This value is used in all presented numerical results, and it has shown a robustness to IF form, SNR level or TF plane dimensions N and K, as it is shown in Section 4.4.
5. The last parameters are the design parameters β and δ. The parameter β is well discussed in wide ACO literature [45]-[59]. It controls the influence of the pheromone on the transition of ants in (13), whereas δ controls the ant's sensitivity to the pheromone concentration [53]-[56]. We use the values of these parameters $\beta = 3.5$, $\delta = 0.2$ as proposed in [53]-[56].

3.6 The IF estimation from the pheromone map

The pheromone map $\mathbf{\Phi}$ can be understood as a new TF representation robust to high noise influence. Thus, the IF estimation problem can be formulated as:

$$\hat{k} = \arg\max_{k} \Phi(n,k). \tag{27}$$

Median filter could be applied on the IF estimation to connect discontinuities (due to dying of agents) in the IF estimation provided by the pheromone map [14]. Since the variable population approach is used to remove as many TF outliers as possible, the IF estimation is conducted based on the kept points. In a high noise environment, several percents of the TF points corresponding to the auto-terms is expected to be removed due to high noise corruption, and thus, interpolation techniques such as cubic interpolation [60] could also be used for the proper IF estimation in regions with missing auto-term values. Instants with eliminated points are easy to detect, since the estimation (27) is equal to

zero at these instants. When the ending points are missing, their values should be set to the nearest non-zero values, or extrapolation can be used.

4. Numerical study and examples

The presented theoretical concepts and the proposed estimation procedure's performance are verified through a set of numerical examples. The WD is calculated using the Hanning window.

We set $I_{\max} = 100$ and all parameters as in Table 4. Median filter with length 3 is applied in the IF estimation (27) in all examples.

Note also that the true IF is obtained from the corresponding non-noisy WD, or other considered TFR. In this way, other error sources, primarily the estimation bias due to IF nonlinearity, are not taken into account. Since our aim is to improve the estimation robustness in high noise, and not to deal with high-accuracy IF estimation, the true IF is adopted in this manner [6]-[8]. The considered synthetic signals have a unity amplitude. Slow variations of the amplitude do not affect the algorithm's accuracy, i.e., these variations affect the accuracy as the overall SNR affects the accuracy of all algorithms analyzed in this section. The considered real signal has a time-varying amplitude.

4.1 Monocomponent signals

The influence of complex zero-mean AWGN $\varepsilon(n)$ with variance σ^2 and with i.i.d. real and imaginary parts is analyzed in all examples. The SNR of the mono-component FM signal (1) is defined as:

$$\text{SNR [dB]} = 10\log_{10}\left(\frac{A^2}{\sigma^2}\right). \tag{28}$$

***Example* 1:** Consider an FM signal

$$x(n) = \exp\left(j\pi \sin\left(\frac{3.5\pi(n+64)}{128}\right) - j35\left(\frac{n+64}{128}\right)^2 \right) + \varepsilon(n) \tag{29}$$

of length $N_s = 256$. The WD is calculated for the middle $N = 128$ points, the signal values are defined at instants $n \in [-N/2, 3N/2)$, and the window length is $K = 128$. The signal is corrupted with noise with SNR = –2dB. The WD of the signal, as the starting point of the proposed algorithm, is shown in Fig. 4 (a). The pheromone maps observed after first 7 iterations, and after the 100-th iteration, are shown in Fig. 4 (b)-(i). It can be observed that after the first iteration a large number of ants have deposited the pheromone at many TF points. However, due to the gradient influence in (14), the pheromone has already started to concentrate at the auto-term, as can be seen from Fig. 4 (b).

At the end of iteration 3 (Fig. 4 (d)), the number of ants is reduced at points outside the auto-term, due to the energy update mechanism (21)-(24), and small gradient (16) values, used in (21).

Figure 5 presents the true IF (circles), the IF estimation obtained using the proposed approach (red solid line), the WD-maxima based estimator (dotted line) and the VA-based estimator (blue solid line) [8].

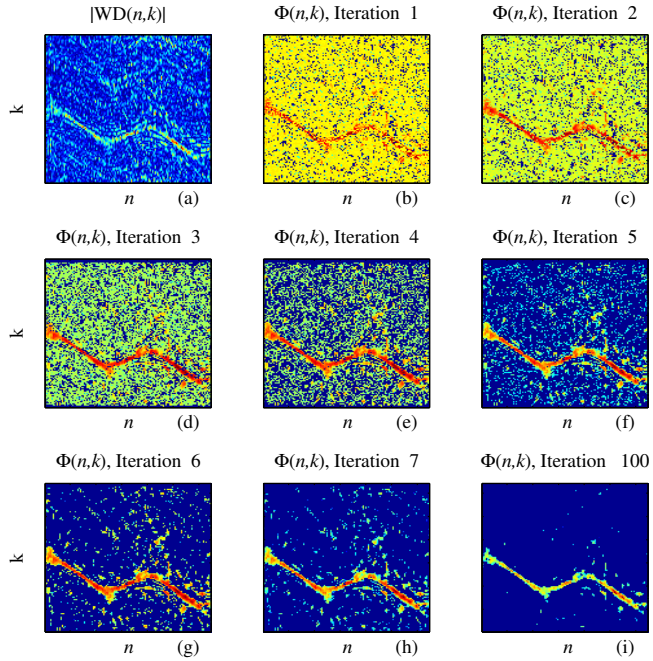

Fig. 4: (a) WD of an FM signal from Example 1; (b)-(i) pheromone map at the end of iterations 1-7 and 100 (last iteration).

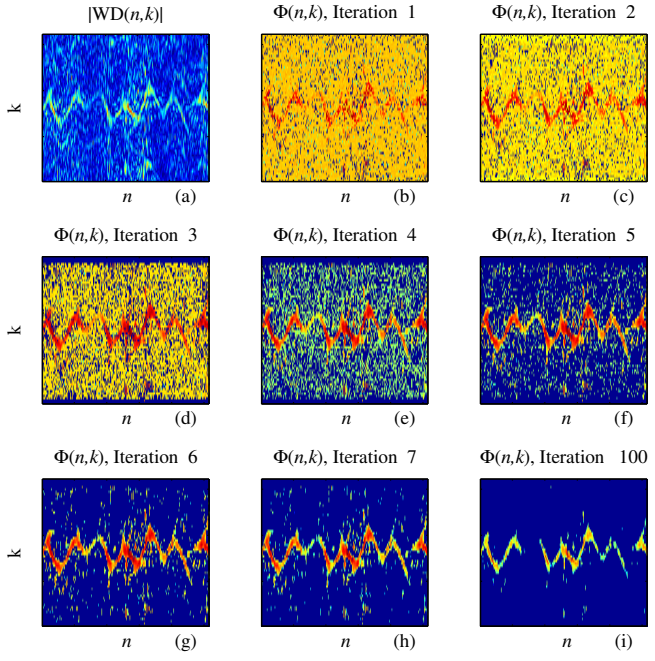

Fig. 6: (a) WD of a FM signal with poly-harmonic modulation from Example 2; (b)-(i) pheromone map at the end of iterations 1-7 and 100 (last iteration).

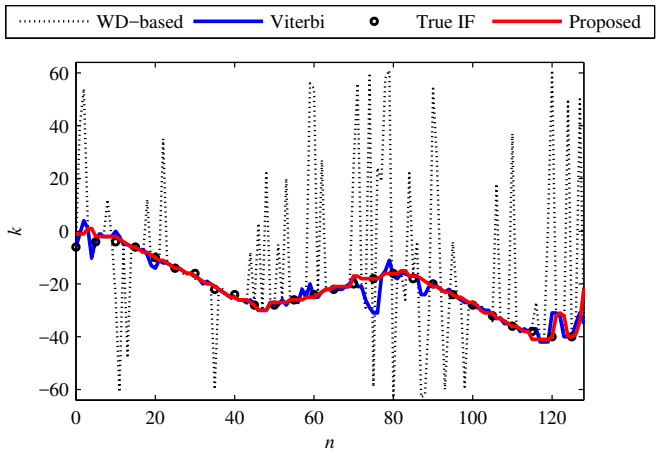

Fig. 5: IF estimation for the signal from Example 1.

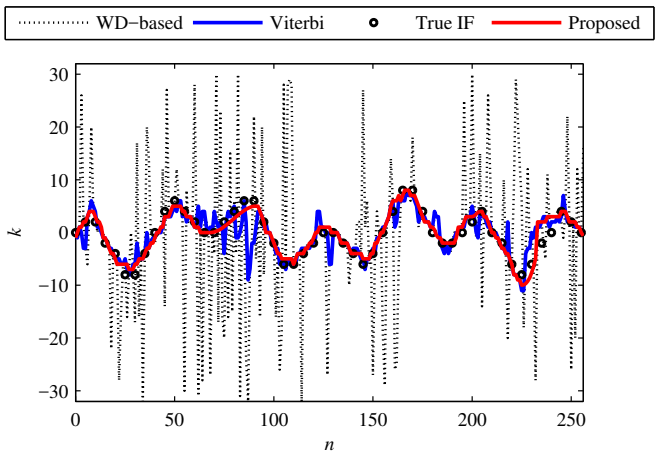

Fig. 7: IF estimation for the signal from Example 2.

***Example* 2**: Here, we consider again a noisy FM signal with poly-harmonic frequency modulation

$$x(n)=\exp\left(j2\sin\left(\frac{13\pi(n+32)}{256}\right)+j3\sin\left(\frac{5\pi(n+32)}{256}\right)\right)+\varepsilon(n) \tag{30}$$

of length N_s = 320, defined at time instants $n\in[-N/2,3N/2)$, where the WD is calculated for the middle N = 256 points. The WD window length is K = 64 and SNR = -2dB.

The pheromone map changes over iterations are shown in Fig. 6 (b)-(i). Figure 7 depicts the true IF, as well as the IF estimations obtained using the proposed method, the WD-maxima approach and VA-based estimator. The proposed approach significantly outperforms the other methods.

Observe the pheromone map (TF representation) after the last iteration (Fig. 6 (i)). It can be noticed that some points corresponding to the IF are missing. However, they are easily detectable and, after the cubic interpolation [60], the IF is reconstructed based on the existing neighboring values.

4.2 Multicomponent signals

The carried out numerical experiments show that the proposed method significantly suppresses WD inner interferences. In order to further illustrate this property, as well as the property of the proposed method to suppress cross-terms, consider a multicomponent signal comprised of two sinusoidally FM signals:

$$x(n)=\exp\left(-j2\pi\sin\left(\frac{-3\pi n}{128}\right)+j\frac{100n}{128}\right)+\exp\left(j2\pi\cos\left(\frac{2\pi n}{128}\right)+j\frac{30\pi n}{128}\right)+\varepsilon(n). \tag{31}$$

In the non-noisy signal case, the cross terms as well as inner interferences are suppressed by the proposed algorithm (Fig. 8, first two plots). It can be explained by the similarity of these disturbances with noisy TF regions, which are detected and suppressed by the proposed pheromone deposition gradient. When the signal is corrupted with noise, the algorithm performs similarly to the non-noisy case. The second two plots in Fig. 8 show the WD and pheromone map of a noisy signal with SNR = -1dB.

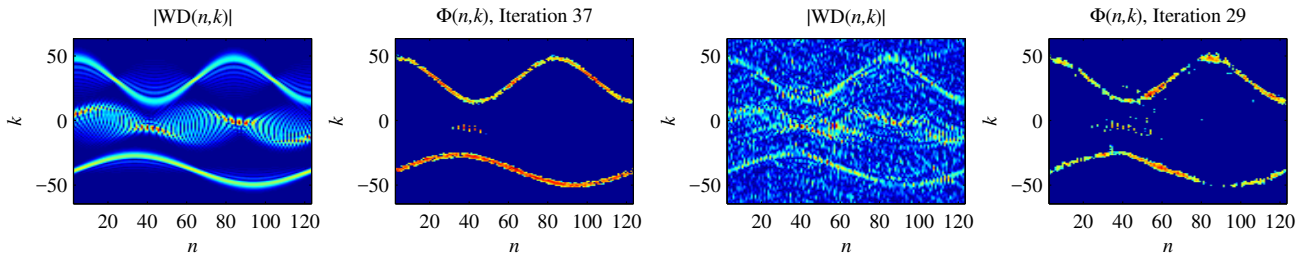

Fig. 8: Suppression of inner interferences and cross-terms reduction in multi-component signals. Noise-free multicomponent signal (first two subplots); Noisy multi-component signal (second two subplots).

The IF estimation of multicomponent signals can be performed using the obtained pheromone map and classical estimation approaches [5], [8], [25]. The approach based on the elimination of the strongest components cannot be successfully applied here, since the pheromone level is not dependent on components amplitudes, and this information is lost in the pheromone map. Hence, the estimation can be performed by identifying the regions of components and estimating IF of each region separately [25].

4.3 Statistical analysis

For all following experiments, we have calculated the MSE versus variable SNR (from -10 to 5dB in Experiment 1 below, and from -10 to 13dB for Experiments 2 and 3 below, with an 1dB step). The MSE is defined as

$$\mathrm{MSE} = E\left[\left|\hat{k}(n) - k_T(n)\right|^2\right], \tag{32}$$

where $\hat{k}(n)$ represents the estimated IF and $k_T(n)$ the true IF frequency index, i.e. position of the WD frequency bin closest to the true frequency for a given n. The MSE is calculated over 100 iterations. We consider the IF estimation using the WD-maxima based approach, the proposed ACO-based algorithm and the VA-based estimator [8]. The IF variations are measured using the instantaneous chirp rate, that is, the second derivative of the signal (1):

$$\mathcal{M}(t) = \phi''(t). \tag{33}$$

For all considered signals, we calculate (33) for their continuous-time form, and then take the maximal absolute value of the discretized instantaneous chirp rate in the considered interval of discrete-time index n for which the WD was calculated.

Experiment 1. Let us consider three signals of the form

$$x(n) = \exp\left(ja\sin\left(\frac{b\pi(n+32)}{256}\right) + jc\sin\left(\frac{d\pi(n+32)}{256}\right)\right) + \varepsilon(n). \tag{34}$$

The signals are of length $N_s = 320$, defined at $n \in [-N/2, 3N/2)$, where the WD is calculated for middle $N = 256$ points. The WD window length is $K = 64$. Parameters defining the three considered signals are given in Table 5.

Table 5

Parameter values used in (34) to define three signals from Experiment 1.

	a	b	c	d	Measure of IF variations based on (33)
Signal 1	1	15	2	5	20.6981
Signal 2	2	13	3	5	31.7746
Signal 3	1	20	1	6	33.5403

We consider the IF estimation using the WD-maxima based approach, the proposed ACO-based algorithm and the VA-based estimator [8]. The obtained MSEs are shown in Fig. 9 (left, middle, right) for signals 1-3 defined as (34) with parameters given in Table 5, respectively. Signals with fast IF variations have been selected in order to confirm good performance of the proposed approach in such scenarios.

Due to the distributed nature of the optimization, the proposed approach is robust to the IF's shape, i.e. to the IF's non-linearity. Observing the results in Fig. 9, we conclude that by slightly increasing the IF variations, similar improvement level of the proposed method over VA-based and the standard WD-based estimator is obtained.

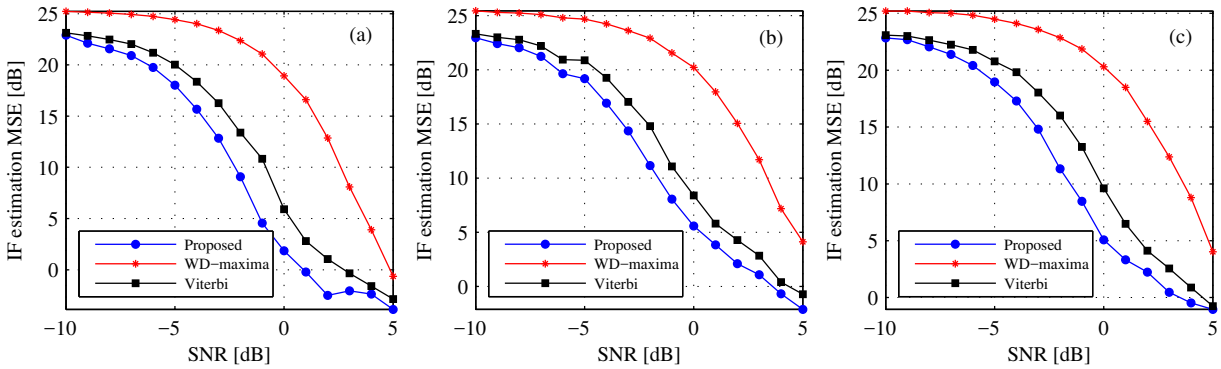

Fig. 9: IF estimation MSE for signals (34) in Experiment 1, using: WD-maxima based estimation (red stars), proposed approach (blue circles) and VA-based approach (black squares).

Experiment 2. Let us now consider four signals of the general form

$$s(n)=\exp\left(ja\sin\left(\frac{b\pi(n+32)}{2M}\right)+jc\cos\left(\frac{d\pi(n+32)}{2M}\right)+je\cos\left(\frac{f\pi(n+32)}{2M}\right)\right)+\varepsilon(n), \tag{35}$$

of length $N_s = 320$, defined at time instants $n\in[-N/2,3N/2)$, where the WD is calculated for the middle $N = 256$ points. The WD window length is $K = 64$. This type of signal was used in [35] to evaluate the level of IF variations. The parameter values in (35) defining considered signals along with the IF variations measure values are given in Table 6.

Table 6

Parameter values used in (35) to define four signals from Experiment 2.

	a	b	c	d	e	f	Measure of IF variations based on (33)
Signal 1	4	27	3	11	10	9	71.5671
Signal 2	3	35	5	10	5	10	90.1180
Signal 3	5	32	10	3	10	3	102.0618
Signal 4	4	41	3	10	5	10	143.2051

The considered signals have different IF shapes, as well as different levels of IF variations. The estimation results are shown in Fig. 10 (a)-(d), for signals 1-4 respectively. The proposed approach outperforms both the WD-maxima based and the VA estimator. It can be observed also that, in general, as the level of IF variations increases, the relative improvement of the proposed method compared with the VA estimator also becomes higher.

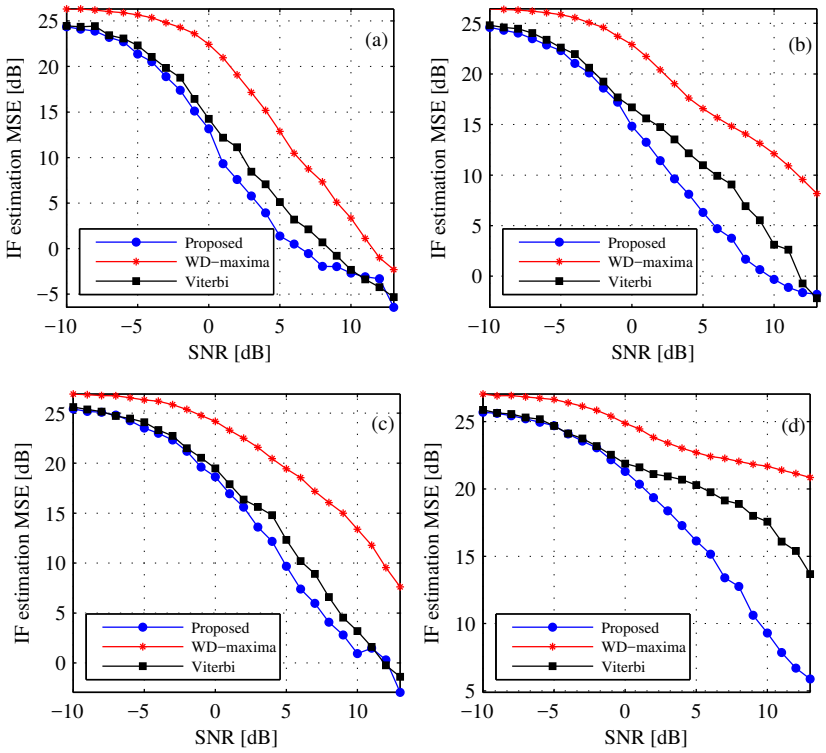

Fig. 10: IF estimation MSE for signals (35) from Experiment 2 using: WD-maxima based estimation (red stars), proposed approach (blue circles) and VA-based approach (black squares).

Experiment 3. With the aim to analyze the influence of the level of IF variations, we consider the signal of the form

$$x(n) = \exp\left(\frac{j30}{a}\sin\left(\frac{a\pi(n+16)}{M}\right)\right) + \varepsilon(n) \tag{36}$$

with length N_s = 160, defined at time instants $n \in [-N/2, 3N/2)$, where the WD is calculated for the middle N = 128 points. The WD window length is K = 32. It can be easily seen that the second derivative of the signal (36) is directly proportional to the constant a. Hence, by increasing a, the introduced measure of IF variations linearly increases. The parameter a is varied in range [13, 21] with step 1. Observing the results in Fig. 11, we conclude that by increasing the IF variations, better performance of the proposed method over the VA-based and the standard WD-based estimator is obtained. Note that the WD-based IF estimation can be affected by fast IF variations even in noiseless case, which in turn can compromise our "True IF" estimation. However, we have kept the value of a within the range that ensures a

meaningful "True IF" estimation. In this paper, the estimation bias due to IF nonlinearity is not taken into account since our aim is to improve the estimation robustness in high noise, and not to deal with bias removal in the WD-based IF estimation [6]-[8].

4.4 Numerical study of parameters

The parameter setup has been determined based on a large number of statistical tests as well as on the results presented in the literature [49]-[56]. Some statistical results for parameter selection are presented next.

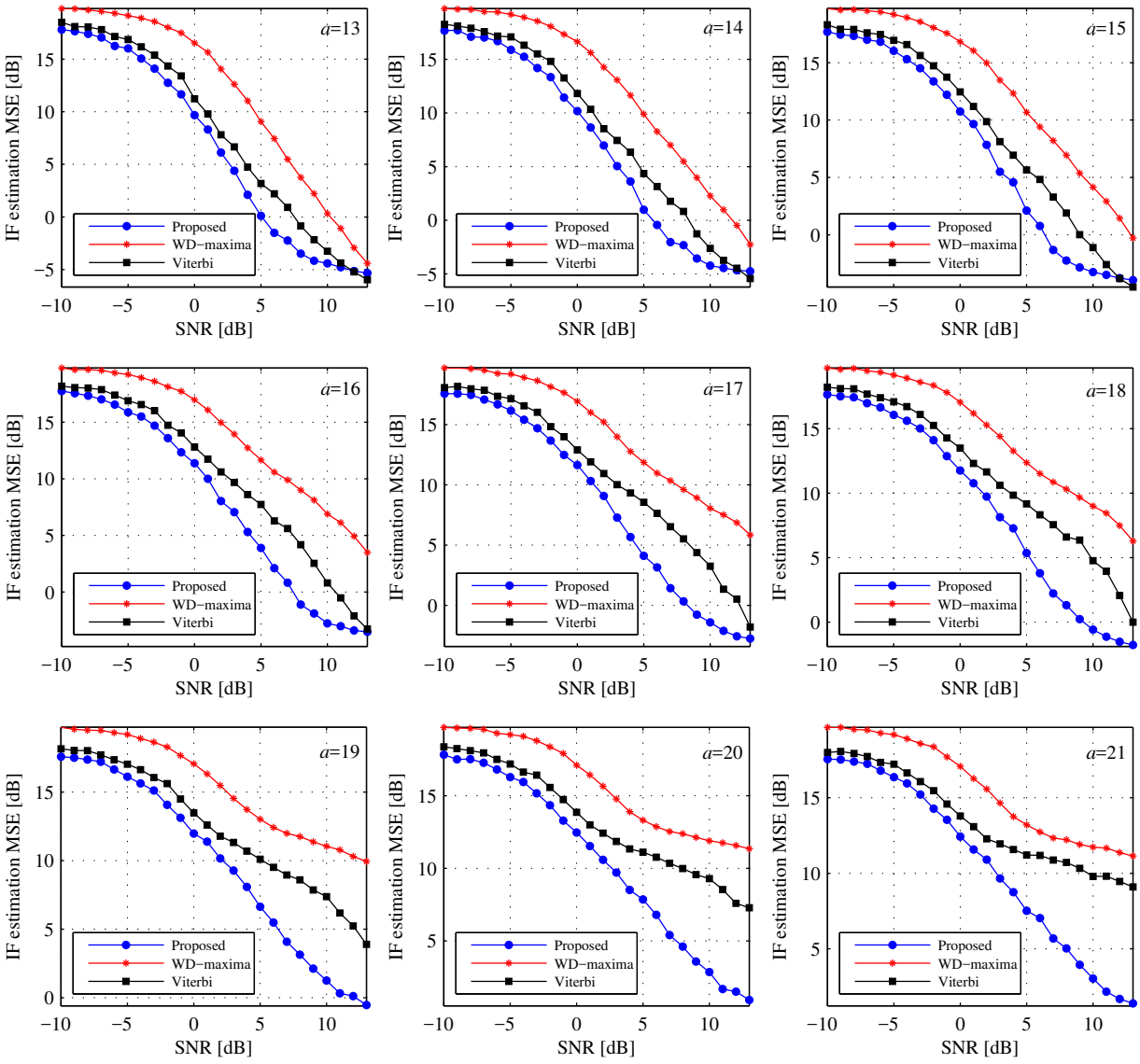

Fig. 11: IF estimation MSE for signal (36) with various *a* values using: WD-maxima based estimation (red stars), proposed ACO-based approach (blue circles) and VA-based approach (black squares).

In our numerical analysis we have observed four noisy signals of the form

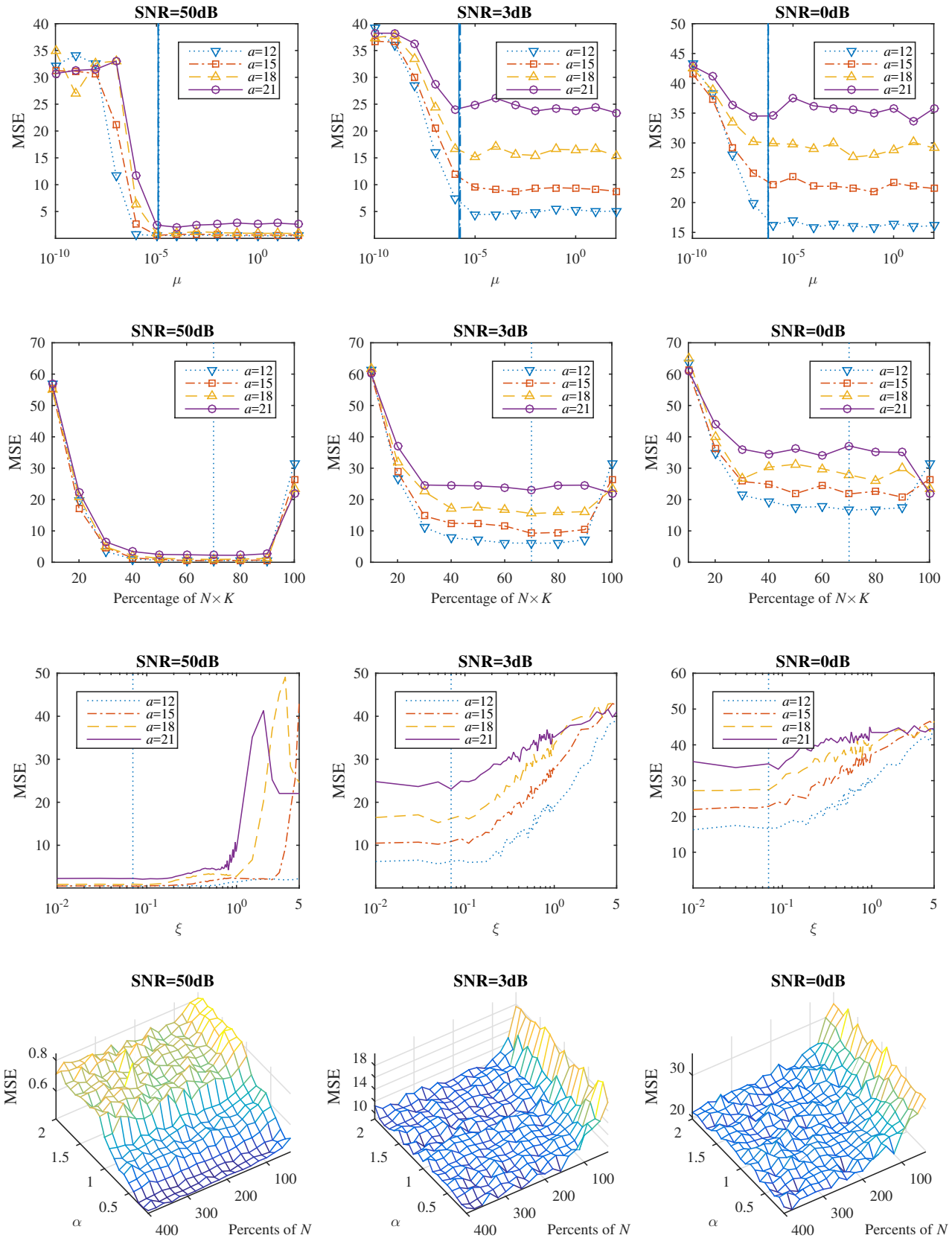

Fig. 12: MSE versus: step μ (first row), initial population size constant γ (second row), pheromone evaporation constant ξ (third row), energy update parameter α and the percent of N used as the stopping criterion (fourth row), shown for three considered SNR levels.

$$x(n) = \exp\left(j\frac{40}{a}\sin\left(\frac{a\pi(n+16)}{256}\right)\right) + \varepsilon(n), \tag{37}$$

with $a \in \{12, 15, 18, 21\}$, having a different level of IF variations, as discussed in Section 4.3. The signal length is N_s = 160, defined at time instants $n \in [-N/2 - K/2,\ N/2 + K/2)$, where the WD is calculated for middle N = 128 points. The WD window length is K = 32. Three different SNR values are considered: 50dB, 3dB, and 0dB. In every experiment, for each observed signal, each SNR level and each parameter value, MSE (32) was calculated based on 100 independent realizations of noisy signals. The last set of parameters was tested on a signal with a = 15.

During the variations of each considered parameter, all remaining parameters were fixed to their values presented in Table 4 (first row).

The step μ was varied in a wide range, from 10^{-10} to 10^{2}, with each value 10 times higher than the previous. The results are shown in Fig. 12 (first row). The vertical line shows the value $\mu = 1/L^4$ used for each considered signal, with L defined in (25), indicating that the maximal WD value can be used to define the lower bound for the step. It can be observed that for each SNR value and each considered signal, the step satisfying (26) produces approximately the same MSE value. It can be also concluded that the estimation breakdown occurs only when a too small step is used.

The initial population size, defined as the percentage γ of the product $N \times K$, was varied from 10 to 100%. The results shown in Fig. 12 (second row) indicate that γ values greater than 50% and smaller than 90% guarantee the lowest MSE. The vertical line indicates the value of 70% used for the results presented in this paper.

The pheromone evaporation rate ξ was varied in a wide range, from 10^{-2} to 1 with the step 0.02 and from 1.5 to 5 with step 0.5. According to the obtained results shown in Fig. 12 (third row) best estimation results are obtained for $\xi < 0.1$. This is in accordance with the results presented in the literature [49], [53], [54], where the value ξ = 0.07 is suggested, which is also used in our numerical results (vertical line).

The energy update constant α was studied together with the stopping criterion of the algorithm, as these two parameters are closely related. Namely, we propose to stop the algorithm when the number of agents falls below 100% of N. Parameter α was varied from 0.1 to 2, whereas the stopping criterion was varied from 40% of N to 400% of N. From Fig. 12 (fourth row) we can observe that the algorithm is quite insensitive to the values of α in the considered range (except for small MSE increase for values larger than 0.5 at high SNR (please have in mind that MSE is not shown in dB)), as the gradient used for the energy update (21) is already scaled with step μ. In this paper, we use α = 0.7.

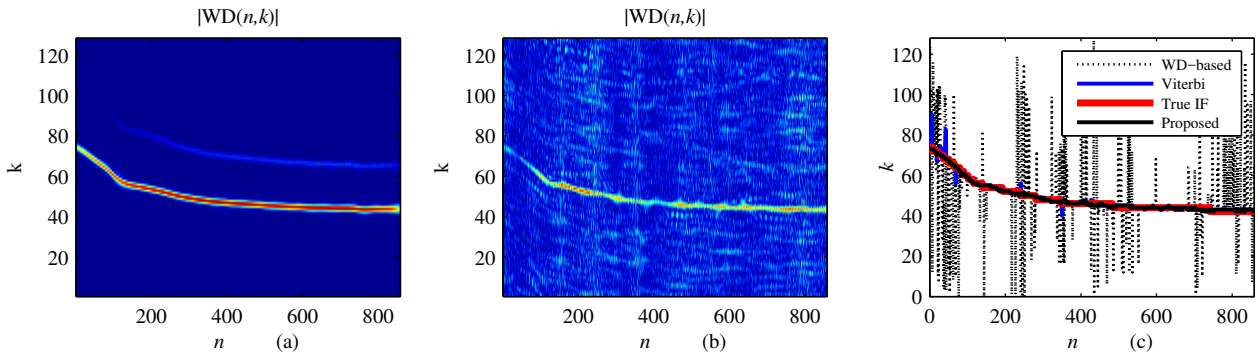

Fig. 13: IF estimation of a bat echolocation sonar signal: (a) WD of the original signal, (b) WD of the signal embedded in high noise and (c) IF estimation results.

4.5 Real signal

We consider a real-life signal to additionally verify the proposed method. A mono-component bat (*Myotis Daubentoni*) echolocation sonar signal $s(n)$ is observed. It was recorded with a sampling frequency of 230.4 kHz and an effective bandwidth of [8 kHz, 80 kHz] as described in [42]. Estimation of parameters of these signals, especially IF, is a well-known problem in TF signal analysis [41]-[43]. We added artificial additive Gaussian noise with SNR=-3dB. Since the observed signal is real-valued, the corresponding Hilbert transform is calculated prior to WD calculation, to avoid the appearance of the component at negative frequencies and cross-terms [6], [9], [28]. The WD of the original signal $s(t)$ and the corresponding noisy version, calculated with the Hanning window of length $K = 128$ are shown in Fig. 13 (a) and (b) respectively. The estimation results are shown in Fig. 13 (c). It can be observed that the WD-maxima based estimator does not provide accurate results except in short ranges where the auto-term is very strong. The VA-based estimator provides accurate estimation results, with the exception of certain number of errors in the first half of the considered time interval where the signal strength is low. As it can be observed in Fig. 13 (c), the IF estimate obtained by the proposed approach matches the true IF.

4.6 Applicability to other time-frequency representations

The presented algorithm can be applied to other TFRs having the property to concentrate the signal energy at or around the IF. The approach has been successfully applied to STFT, S-method [29], reassigned spectrogram [36], [37], and wavelet transform [30], [32], [39]. TFRs are shown for clean and noisy signals, as well as the pheromone maps obtained in the noisy signal cases. Estimation results are shown for TFR maxima approach (in case of complex TFRs absolute value is exploited) and for the proposed technique and the VA applied on noisy TFRs.

The noisy signal of the form

$$x(n)=\exp\left(j22\sin\left(\frac{14\pi(n+32)}{2M}\right)+j16\cos\left(\frac{7\pi(n+32)}{2M}\right)+j16\cos\left(\frac{5\pi(n+32)}{2M}\right)\right)+\varepsilon(n) \quad (38)$$

is considered for the case of STFT (SNR level of -1dB), whereas the signal

$$x(n)=\exp\left(j8\sin\left(\frac{36\pi(n+32)}{2M}\right)+j6\cos\left(\frac{9\pi(n+32)}{2M}\right)+j10\cos\left(\frac{8\pi(n+32)}{2M}\right)\right)+\varepsilon(n) \quad (39)$$

is considered in the case S-method (SNR level 4dB). As the STFT is complex valued, its absolute value is used as the ACO algorithm input. Both signals have length N_s = 320, and they are defined at time instants $n\in[-N/2,3N/2)$, where the TFR is calculated for the middle N = 256 points, with window length K = 64.

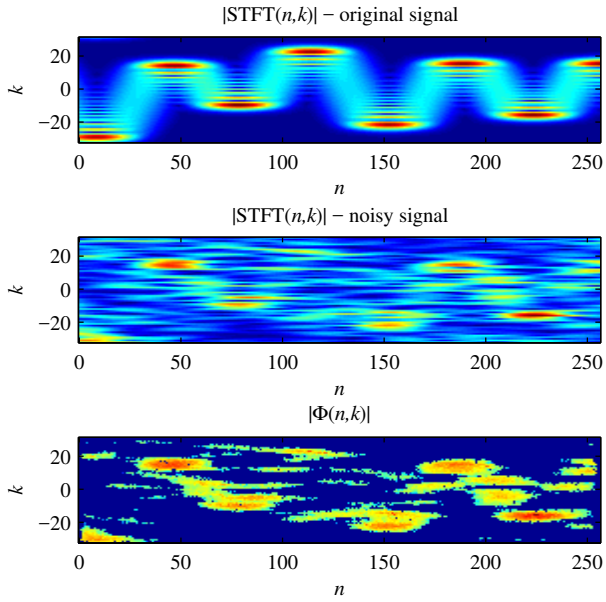

Fig. 14: STFT of a non-noisy sinusoidally modulated signal (first row) and the noisy signal (second row); the obtained pheromone map (third row).

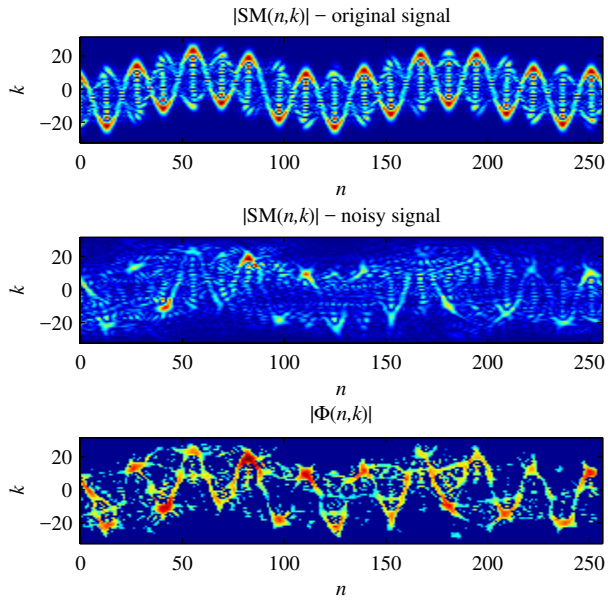

Fig. 16: S-method of a non-noisy sinusoidally modulated signal (first row and the noisy signal (second row); the obtained pheromone map (third row).

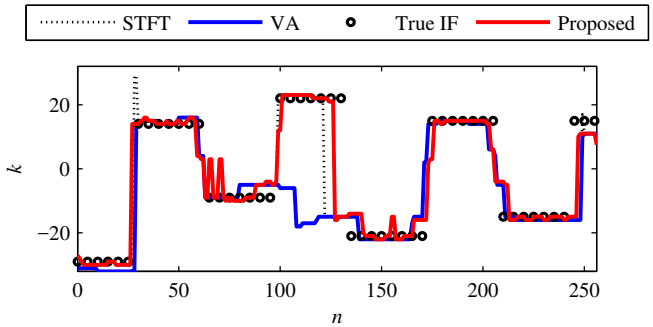

Fig. 15: IF estimation results for the signal (38) based on STFT

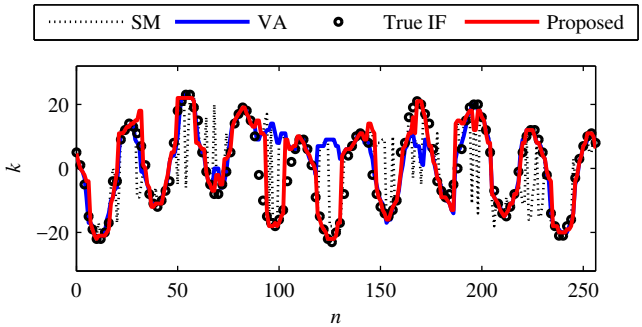

Fig. 17: IF estimation results for the signal (39) based on S-method

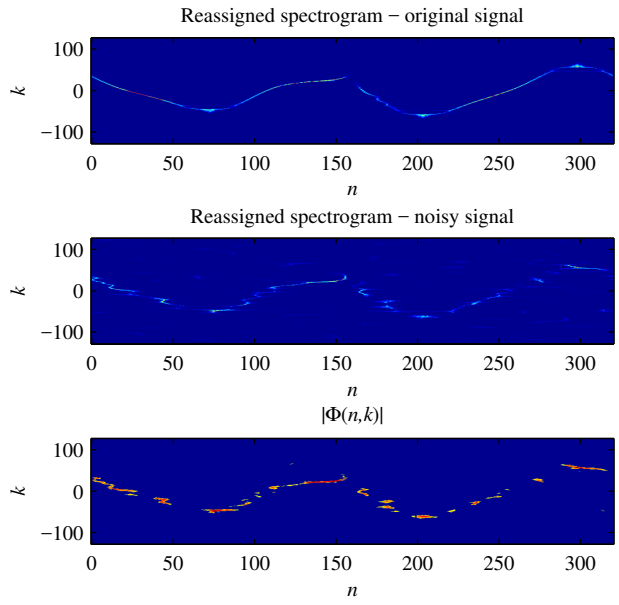

Fig. 18: Reassigned spectrogram of a non-noisy sinusoidally modulated signal (first row) and the noisy signal (second row); the obtained pheromone map (third row).

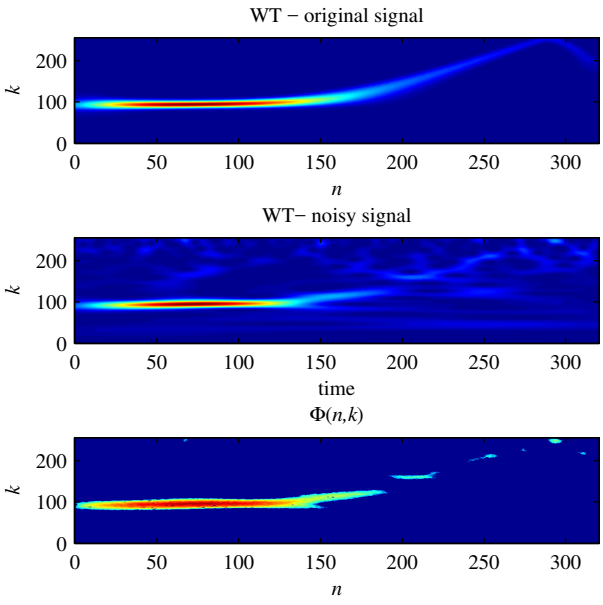

Fig. 20: Scalogram of a non-noisy sinusoidally modulated signal (first row) and the noisy signal (second row); the obtained pheromone map (third row).

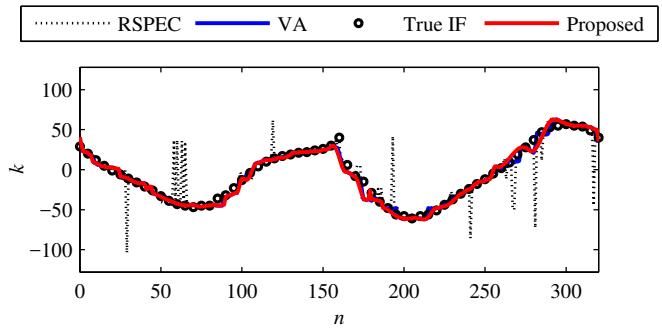

Fig. 19: IF estimation results for the signal (40) based on the reassigned spectrogram

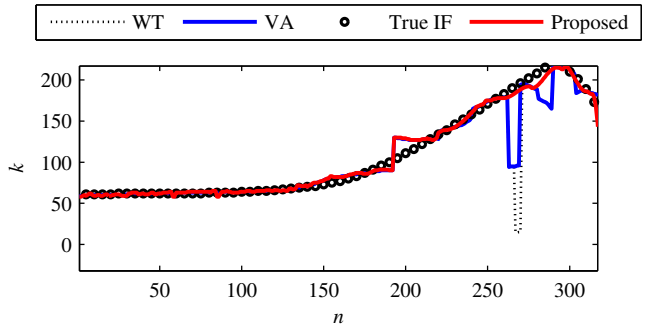

Fig. 21: IF estimation results for the signal (41) based on the wavelet transform

The STFT of the non-noisy and noisy signal as well as the obtained pheromone map are shown in Fig. 14. The estimation results are given in Fig.15 for the proposed method, STFT absolute value maxima and the corresponding VA estimator. For the case of the S-method, the results are shown in Figs. 16 and 17.

For the case of reassigned spectrogram, we consider the signal

$$x(n)=\exp\left(j2\sin\left(\frac{15\pi(n+32)}{2M}\right)+j22\cos\left(\frac{6\pi(n+32)}{2M}\right)+j20\cos\left(\frac{5\pi(n+32)}{2M}\right)\right)+\varepsilon(n)\ . \qquad (40)$$

The signal length is N_s = 320. The reassigned spectrogram and the obtained pheromone map are shown in Fig. 18, where the SNR of 1dB is considered in the noisy signal. From Fig. 19 we can observe that the reassigned spectrogram maxima based estimator has outliers that both ACO and VA estimators can correct.

In the case of wavelet transform (Morlet wavelet) the PPS signal

$$x(n)=\exp\left(j5\pi\left(\frac{n+32}{2M}\right)^{7}-j50\left(\frac{n+32}{2M}\right)\right)+\varepsilon(n) \tag{41}$$

is considered. The signal length is Ns = 320. The scalogram is shown with discrete frequency indices k in Fig. 20, whereas the estimation results for wavelet transform is shown in Fig. 21 for the SNR level of -4dB. We may observe that the proposed approach shows better results at higher frequencies where the signal energy concentration is low.

4.7 Time-frequency plane splitting on non-overlapped subintervals

The PPS signal

$$x(n)=\exp\left(-j1200\pi\left(\frac{n-96}{512}\right)^{3}+j100\pi\left(\frac{n+32}{512}\right)\right)+\varepsilon(n) \tag{42}$$

of length N_s = 320, defined at time instants $n\in[-N/2,\,3N/2)$ is considered, where the WD is calculated for the middle N = 256 points. The WD window length is K = 64. The WD TFR plane is divided into non-overlapped strips of width 8, 16, 32, and 64. The proposed algorithm is applied on each strip (subinterval of the time axis) separately and the MSE in the IF estimation was calculated, for various SNR levels based on 100 signal realizations for each SNR level. The results are provided in Table 7. For comparison, the MSE results obtained when the algorithm is applied on the entire TF plane are also presented in the table. We can conclude that good results are obtained even for subintervals of length 16 and 32, whereas in the case when the TFR plane is divided in strips of width 64 results are very close to those obtained when the algorithm is applied on the entire TF plane.

Table 7
MSE of the IF estimation based on WD and the proposed algorithm shown for different subinterval lengths

SNR [dB]	MSE [dB] subinterval size: 8	MSE [dB] subinterval size: 16	MSE [dB] subinterval size: 32	MSE [dB] subinterval size: 64	MSE [dB] Entire TF plane
3	12.69	-2.29	-4.79	-4.95	-5.07
2	12.22	-2.54	-4.06	-4.81	-5.06
1	12.71	1.31	-1.53	-3.10	-3.54
0	12.63	5.41	-0.08	-0.12	-1.06
-1	14.68	8.88	7.99	4.06	3.64
-2	16.55	13.44	9.99	10.53	10.52
-3	17.98	16.48	15.06	14.70	13.64

5. Conclusion

The modified ACO-based algorithm is proposed for the IF estimation in high Gaussian noise environments. The algorithm is adapted to the IF estimation from the WD of the considered signal. The pheromone deposition gradient is developed based on the two important IF properties: IF should pass through as many as possible points of the WD with highest magnitudes, whereas the IF variations between two consecutive points should not be too fast. The varying population concept has been introduced to improve the IF estimation. The estimation performance of the proposed algorithm is illustrated with numerical experiments including a real-life signal, which clearly confirm that it outperforms the WD-maxima based estimator in high noise environments. Moreover, it outperforms the state-of-the-art VA-based estimator when IF fast variations are involved. The presented approach is applicable to other time-frequency representations as well. Algorithm parameters were studied in extensive statistical test, showing the robustness of the approach to parameter selection in wide ranges. The algorithm was compared with the VA and the WD-maxima estimators in statistical tests with various signals with different levels of IF variations, measured using the maximal instantaneous chirp rate of the considered signals. The presented algorithm can be applied in post-processing of WD, reducing inner interferences and cross-components in multi-component signals. The applications in post-processing of multi-component signals with close components as well as in real practical scenarios are the part of our further research.

Acknowledgment

This work has been supported by the HERIC project through the BIO-ICT Centre of Excellence (Contract No. 01-1001) and the Ministry of Science of Montenegro.

Appendix A.

Summarization of the basic ACO algorithm is given as follows.

***Step* 0**: Place the agents at random positions and with random orientations over the discrete grid, by forming the matrix **P** according to (8). Initialize the pheromone matrix **Φ** according to (9).

Repeat steps 1-3 I_{max} times:

***Step* 1**: For every agent compute the probability (13), and move the agent to an adjacent cell which

- is characterized by the highest probability and

- is not occupied by other ants.

Update the corresponding elements of the matrix **P.**

***Step* 2**: For each grid point (*n, k*) visited by an agent in Step 1 update the matrix **Φ** using (14).

***Step* 3**: Update the pheromone matrix **Φ** according to (15).

Appendix B.

The proposed algorithm for the pheromone map generation can be summarized as follows.

***Input**: time-frequency matrix* of size $N \times K$

Initialization:

Place ants at initial random positions (n_0, k_0), by forming the auxiliary matrix **P** according to (8) and initialize the pheromone matrix **Φ** and the energy matrix **E** according to (9) and (20), respectively. Form the auxiliary orientation matrix **d** according to (11).

***Calculations*:**

While the number of iterations is less than $I_{\max}$ and the number of non-zero elements in **P** is larger than 80-100% of *N* repeat:

Step 1: For every non-zero element in **P** compute probabilities (13), based on (11) and (12), and move the agent to an adjacent cell which is characterized by the highest probability and which is not occupied by other ants.

Update the corresponding elements of position matrix **P**.

Step 2: For every grid point visited by the agent in Step 1, calculate the gradient ∇(*n*, *k*) according to (16). Update the pheromone matrix **Φ** using (14) and the energy matrix **E** using (21) and (22).

Step 3: Update the pheromone matrix **Φ** according to (15). Update the energy matrix **E** according to (23). Update the matrix **P** according to (24).

***Output*:** *pheromone matrix* **Φ**

References

[1] B.Boashash (Ed.), *Time-Frequency Signal Analysis and Processing, A Comprehensive Reference*, Elsevier, The Boulevard, Langford Lane, Kidlington, Oxford, UK, 2003

[2] B. Boashash, P. O' Shea, "Use of the cross Wigner-Ville distribution for estimation of instantaneous frequency," *IEEE Transactions on Signal Processing*, Vol. 41, No.3, pp.1439-1445, March 1993.

[3] L. Cohen, *Time–Frequency Analysis*, Prentice-Hall, New York, 1995.

[4] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal –Part 1: Fundamentals," *Proceedings of the IEEE*, Vol. 80, No. 4, pp.519-538, April 1992.

[5] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal: A tutorial review-Part 2: algorithms and applications," *Proceedings of the IEEE*, Vol. 80, No. 4, pp.540-568, April 1992.

[6] Lj. Stanković, I. Djurović, S. Stanković, M. Simeunović, S. Djukanović, and M. Daković, "Instantaneous frequency in time–frequency analysis: Enhanced concepts and performance of estimation algorithms," *Digital Signal Processing*, Vol. 35, pp.1-13, December 2014.

[7] I. Djurović, Lj. Stanković, "Influence of high noise on the instantaneous frequency estimation using time-frequency distributions," *IEEE Signal Processing Letters*, Vol. 7, No.11, pp.317 – 319, November 2000.

[8] I. Djurović, Lj. Stanković, "An algorithm for the Wigner distribution based instantaneous frequency estimation in a high noise environment," *Signal Processing*, Vol. 84, No. 3, pp.631-643, March 2004.

[9] I. Djurović, "Viterbi algorithm for chirp-rate and instantaneous frequency estimation," *Signal Processing*, Vol. 91, No. 5, pp.1308-1314, May 2011.

[10] Z. Xiaoping; Z. Xiuli; Z. Yuhui; X. Shengdong,"A robust instantaneous frequency estimation of rotating machinery," *International Conference on Mechatronic Science, Electric Engineering and Computer (MEC), 2011*, pp.1922-1925, August 2011.

[11] P. Li, D. Wang, L. Wang, "Separation of micro-Doppler signals based on time frequency filter and Viterbi algorithm," *Signal, Image and Video Processing,* Vol. 7, No. 3, pp.593-605, May 2013.

[12] C. Wang, F. Kong, Q. He, F. Hu, F. Liu, "Doppler effect removal based on instantaneous frequency estimation and time domain re-sampling for wayside acoustic defective bearing detector system," *Measurement*, Vol. 50, pp.346-355, April 2014.

[13] C. Ioana, J. Mars, A. Serbanescu, S. Stanković, "Time-frequency-phase tracking approach: application to underwater signals in a passive context," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2010)*, pp.5634-5637, March 2010.

[14] I. Djurović, Lj. Stanković, "Modification of the ICI rule based IF estimator for high noise environments," *IEEE Transactions on Signal Processing*, Vol. 52, No. 9, pp. 2655-2661, September 2004.

[15] J. Lerga, V. Sucic, “Nonlinear IF estimation based on the pseudo WVD adapted using the improved sliding pairwise ICI rule,” *IEEE Signal Processing Letters*, Vol. 16, No. 11, pp. 953–956, Nov. 2009.

[16] V. C. Chen, H. Ling, *Time-frequency transforms for radar imaging and signal analysis*, Artech House, Boston, 2002.

[17] S. Marchand, “The Simplest Analysis Method for Non-Stationary Sinusoidal Modeling,” *Proceedings of the Digital Audio Effects (DAFx'12) Conference*, pp. 23-26, York, United Kingdom, September 2012.

[18] F. Auger, P. Flandrin, Y-T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, H-T. Wu, “Time-frequency reassignment and synchrosqueezing : An overview”, *IEEE Signal Processing Magazine*, Vol. 30, No. 6, pp. 32-41, 2013.

[19] T. Oberlin, S. Meignen, S. McLaughlin, “A novel Time-Frequency technique for multicomponent signal denoising,“ *European Signal Processing Conference (EUSIPCO)*, Marrakech, Morocco, Sept. 9-13, 2013.

[20] D. Iatsenko, P.V.E. McClintock, A. Stefanovska, “Extraction of instantaneous frequencies from ridges in time–frequency representations of signals”, *Signal Processing*, Vol. 126, pp. 290–303, August 2016

[21] M. Jabloun, N. Martin, F. Leonard, M. Vieira, "Estimation of the instantaneous amplitude and frequency of non-stationary short-time signals." *Signal processing*, Vol 88., No. 7, pp. 1636-1655, July 2008

[22] M. Jabloun, F. Leonard, M. Vieira, N. Martin, "A New Flexible Approach to Estimate the IA and IF of Nonstationary Signals of Long-Time Duration," *IEEE Transactions on Signal Processing*, Vol. 55, No. 7, pp. 3633-3644, July 2007.

[23] M. Aoi, K. Lepage, Y. Lim, U. T. Eden, T. J. Gardner, "An Approach to Time-Frequency Analysis With Ridges of the Continuous Chirplet Transform," *IEEE Transactions on Signal Processing*, vol. 63, no. 3, pp. 699-710, Feb.1, 2015.

[24] C. Wang, M. Amin, "Performance analysis of instantaneous frequency-based interference excision techniques in spread spectrum communications," *IEEE Transactions on Signal Processing*, Vol. 46, No. 1, pp.70-82, January 1998.

[25] Lj. Stanković, M. Daković, T. Thayaparan, *Time-Frequency Signal Analysis with Applications*, Artech house, 2013.

[26] P. Rao, F.J. Taylor, "Estimation of instantaneous frequency using the discrete Wigner distribution," *Electronics Letters*, Vol. 26, No. 4, pp.246-248, February 1990.

[27] V. Katkovnik, Lj. Stanković, "Instantaneous frequency estimation using the Wigner distribution with varying and data-driven window length," *IEEE Transactions on Signal Proccessing,* Vol. 46, No. 9, pp.2315-2326, September 1998.

[28] V. N. Ivanović, M. Daković, Lj. Stanković, "Performance of Quadratic Time-Frequency Distributions as Instantaneous Frequency Estimators," *IEEE Transactions on Signal Processing*, Vol. 51, No. 1, pp.77-89, Jan. 2003.

[29] LJ. Stanković, "A method for time-frequency signal analysis," *IEEE Transactions on Signal Processing*, Vol. 42, No. 1, pp. 225-229, January 1994.

[30] R. A. Carmona, W. L. Hwang, B. Torresani, "Characterization of signals by the ridges of their wavelet transforms," *IEEE Transactions on Signal Processing*, vol.45, no.10, pp.2586-2590, Oct 1997.

[31] R. A. Carmona, W. L. Hwang, B. Torresani, "Multiridge detection and time-frequency reconstruction," *IEEE Transactions on Signal Processing*, Vol.47, No.2, pp.480-492, February 1999.

[32] A. Grossmann, J. Morlet, "Decomposition of Hardy functions into square integrable wavelets of constant shape," *Soc. Int. Am. Math. (SIAM), J. Math. Analys*., Vol. 15, No. 4, pp. 723-736, 1984.

[33] I. Djurović, C. Ioana, T. Thayaparan, LJ. Stanković, P Wang, V. Popović-Bugarin, M. Simeunović, "Cubic-phase function evaluation for multicomponent signals with application to SAR imaging, " *IET Signal Processing*, Vol. 4, No. 4, pp. 371-381 August 2010.

[34] J. Lerga, V. Sucic, B. Boashash, "An Efficient Algorithm for Instantaneous Frequency Estimation of Nonstationary Multicomponent Signals in Low SNR," *EURASIP J. Adv. Signal Process*., 2011, doi:10.1155/2011/725189

[35] A. Omidvarnia, G. Azemi, J. M. O'Toole, B. Boashash, "Robust estimation of highly-varying nonlinear instantaneous frequency of monocomponent signals using a lower-order complex-time distribution," *Signal Processing*, No. 93, pp. 3251-3260, 2013.

[36] F. Auger, P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Transactions on Signal Processing*, Vol. 43, No. 5, pp. 1068–1089., May 1995.

[37] K. Kodera, R. Gendrin, C. de Villedary, "Analysis of time-varying signals with small BT values," *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 26, No. 1, pp. 64–76, Feb. 1978.

[38] K. Czarnecki, "The instantaneous frequency rate spectrogram," *Mechanical Systems and Signal Processing*, Vol. 66, No. 67, pp. 361-373, 2016

[39] J. P. de León, J. R. Beltrán, F. Beltrán, "Instantaneous frequency estimation and representation of the audio signal through Complex Wavelet Additive Synthesis", *International Journal of Wavelets, Multiresolution and Information Processing*, Vol. 12, No. 3, 2014.

[40] V. Katkovnik, I. Djurović, LJ. Stanković, "Robust Time-Frequency Distributions," *in Time-Frequency Signal Analysis and Processing,*, ed. B. Boashash, Academic Press, December 2015.

[41] Y. Kopsinis, E. Aboutanios, D. A. Waters, S. McLaughlin, "Investigation of bat echolocation calls using high resolution spectrogram and instantaneous frequency based analysis," *2009 IEEE/SP 15th Workshop on Statistical Signal Processing*, pp. 557-560, Cardiff, 2009.

[42] P. Flandrin, "On detection-estimation procedures in the time-frequency plane," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '86)*, pp. 2331-2334, 1986.

[43] Y. Yang, W. Zhang, Z. Peng, G. Meng, "Multicomponent Signal Analysis Based on Polynomial Chirplet Transform," *IEEE Transactions on Industrial Electronics*, Vol. 60, No. 9, pp. 3948-3956, September 2013.

[44] E. Bonabeau, M. Dorigo, G. Theraulaz, *Swarm Intelligence: From Natural to Artificial Systems*, Oxford University Press, 1999.

[45] M. Dorigo, T. Stützl, *Ant Colony Optimization*, Bradford Co. Scituate, MA, USA, 2004

[46] M. Dorigo, M. Birattari, T. Stutzle, "Ant colony optimization," *IEEE Computational Intelligence Magazine*, Vol. 1, No. 4, pp. 28-39, November 2006.

[47] M. Dorigo, L.M. Gambardella, "Ant colony system: a cooperative learning approach to the traveling salesman problem," *IEEE Transactions on Evolutionary Computation*, Vol. 1, No.1, pp.53-66, April 1997.

[48] M. Dorigo, V. Maniezzo, A. Colorni, "Ant system: optimization by a colony of cooperating agents," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, Vol. 26, No. 1, pp.29-41, February 1996.

[49] C. Fernandes, V. Ramos, A.C. Rosa, "Self-regulated artificial ant colonies on digital image habitats," *International Journal of Lateral Computing*, Vol. 2, No.1, December 2005.

[50] X. Liu, S. Fang, "A convenient and robust edge detection method based on ant colony optimization," *Optics Communications*, Vol. 353 , pp. 147-157, October 2015.

[51] H. Nezamabadi-pour, S. Saryazdi, E. Rashedi, "Edge detection using ant algorithms," *Soft Computing*, Vol. 10, No. 7, pp. 623–628, August 2005.

[52] B. Chen, L. Chen, Y. Chen, "Efficient ant colony optimization for image feature selection", *Signal Processing*, Vol. 93, No. 6, pp. 1566-1576, June 2013.

[53] C. Fernandes, V. Ramos, A. C. Rosa, "Varying the population size of artificial foraging swarms on time varying landscapes, "*Proc. of Artificial Neural Networks: Biological Inspirations – ICANN 2005,* Vol. 3696, Part I, pp.311-316, September 2005.

[54] V. Ramos, F. Almeida, "Artificial ant colonies in digital image habitats – A mass behaviour effect study on pattern recognition," *Proc. ANTS´2000*, pp.113-116, September 2000.

[55] A. Mora, C. M. Fernandes, J. J. Merelo, V. Ramos, J. L. J. Laredo, A.C. Rosa, "KohonAnts: A Self-Organizing Ant Algorithm for Clustering and Pattern Classification, " *Proceedings of the 11th International Conference on Artificial Life, S. Bullock et al., Eds., MIT Press*, Cambridge, MA, pp. 428-435, 2008.

[56] A. Jevtić, D. Andina, "Adaptive artificial ant colonies for edge detection in digital images," *IECON 2010 - 36th Annual Conference on IEEE Industrial Electronics Society*, pp.2813-2816, November 2010.

[57] S. Le Hegarat-Mascle, A. Kallel, X. Descombes, "Ant Colony Optimization for Image Regularization Based on a Nonstationary Markov Modeling," *IEEE Transactions on Image Processing*, Vol. 16, No. 3, pp. 865-878, March 2007.

[58] GK. Apostolidis, LJ. Hadjileontiadis, "Swarm decomposition: A novel signal analysis using swarm intelligence," *Signal Processing*, Vol.132, 2017.

[59] M. Brajović, V. Popović-Bugarin, “Instantaneous frequency estimation using ant colony optimization and Wigner distribution,” *4th Mediterranean Conference on Embedded Computing - MECO 2015*, pp.349-352, June 2015.

[60] F. N. Fritsch, R. E. Carlson, "Monotone piecewise cubic interpolation," *SIAM Journal on Numerical Analysis (SIAM)*, Vol. 17, No. 2, pp.238-246, 1980.