Immune-Inspired Optimization with Autocorrentropy Function for Blind Inversion of Wiener Systems

Stephanie A. Fernandez*†, Denis G. Fantinato‡, Jugurta Montalvão§, Romis Attux¶ and Daniel G. Silva†
* IN3 – Computer Science, Multimedia and Telecommunication Dept., Open University of Catalonia, Spain
Email: {salvarezferna}@uoc.edu
† Dept. of Electrical Engineering, University of Brasília, Brazil
E-mail: danielgs@ene.unb.br
‡ Mathematics, Computation and Cognition Center, Federal University of ABC, Brazil
E-mail: denis.fantinato@ufabc.edu.br
§ Dept. of Electrical Engineering, University of Sergipe, Brazil
E-mail: jmontalvao@ufs.br
¶ School of Electrical and Computer Engineering, University of Campinas, Brazil
E-mail: attux@dca.fee.unicamp.br

***Abstract*—Blind inversion of nonlinear systems is a complex task that requires some sort of prior information about the source e.g. whether it is composed of independent samples or, particularly in this work, a dependence "signature" which is assumed to be known via the autocorrentropy function. Furthermore, it involves the solution of a nonlinear, multimodal optimization problem to determine the parameters of the inverse model. Thus, we propose a blind method for Wiener systems inversion, which is composed of a correntropy-based criterion in association to the well-known CLONALG immune-inspired optimization metaheuristic. The empirical results validate the methodology for continuous and discrete signals.**

I. Introduction

System Identification is an important task in view of its vast horizon of practical applications. In overall, it is a technique that allows to build, from measured data, mathematical models for dynamic systems. A dynamic system can be analyzed in time and/or in frequency. For this reason, the identification task must be able to derive linear or nonlinear models. Over the years, the study of nonlinear systems has grown quickly with the technological and industry advance. In general, a nonlinear dynamic system shows a complex behavior and its future state can be considerably changed from small variations in the initial conditions [1].

Traditionally, nonlinear system identification methods assume that a reference signal is available. However, in a real-world situation, one may have no access to the system input, hence, the blind inversion of nonlinear systems is a necessary strategy to circumvent such restriction. Blind inversion (closely related to blind identification and equalization) of nonlinear systems has important developments, with many theoretical results and applications, for instance, in telecommunications [2], [3]. In blind inversion problems, each individual input, which is to be measured, goes through the system and the corresponding output is observed. Both input values and the system to be inverted are unknown.

The well-known Wiener and Hammerstein systems are nonlinear models that are employed within many domains, due to their simplicity and physical meaning. The system steady-state behavior is completely determined by static nonlinearities, while the system dynamic behavior is determined by both the nonlinearity and the linear model components. Formally speaking, a Wiener system, depicted in Fig. 1a, consists of a linear time-invariant (LTI) filter subsystem $h(n)$ followed by a memoryless, invertible, nonlinear distortion $f[\cdot]$:

$$x(n) = f\left[e(n)\right] = f\left[h(n) * s(n)\right], \tag{1}$$

where $s(n)$ is the system input signal and $x(n)$ is its output. A Hammerstein system, depicted in Fig. 1b is composed by a static nonlinear block $g[\cdot]$ followed by an LTI subsystem $w(n)$:

$$y(n) = w(n) * u(n) = w(n) * g\left[x(n)\right], \tag{2}$$

where $x(n)$ is the input signal to the Hammerstein system and $y(n)$ its output.

In this work, we consider the inverse modeling of a Wiener system by means of immune-inspired engineering. Despite the system simplicity, it has been applied in many areas, such as industry, sociology, psychology etc. [4], [5].

In the last decade, Information Theoretic Learning (ITL) has gained attention in the signal processing area [6]–[8] and a new generalized correlation function, called correntropy, has been introduced. Correntropy is a positive definite function which yields a generalized similarity measure between random variables (in this work, sets of random variables are used to model signal samples through time, thus yielding stochastic

978-1-5090-6017-7/18/$31.00 ©2018 IEEE

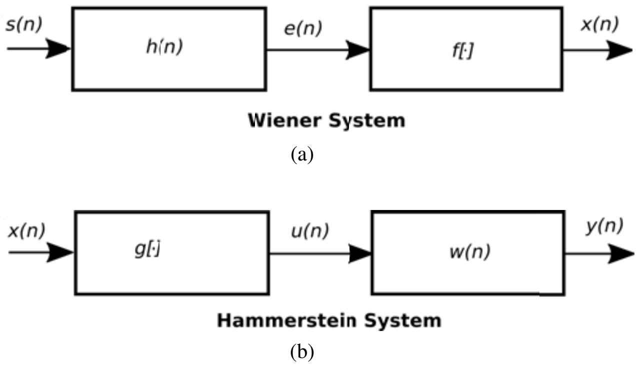

Fig. 1: Diagram of Wiener and Hammerstein systems.

processes) capable of encompassing the higher-order statistics (HOS) of the signals of interest. Recently, the authors presented in [9] a first analysis of the autocorrentropy and autocorrelation function for representing the time structure of a given signal in the context of the unsupervised inversion of Wiener systems by Hammerstein systems. The results indicated that both functions provide effective means for system inversion, being possible to visualize the effect of linear feedback on the overall system performance.

This work is an extension of the previous one, with the aim of (*i*) introducing a new antibody/cell representation, based on the zeros and poles of the inverse filter, to avoid unstable solutions; (*ii*) analyzing the correntropy-based criterion in more scenarios and its sensitivity to the kernel size adjustment; and (*iii*) confirming the adequacy of CLONALG immune-inspired algorithm as optimization strategy for this task.

The paper is divided into 6 sections. In the following, Section II reviews the main publications on the topic. Section III details the criterion to be employed through the remainder of the work. Section IV briefly presents the proposed methodology. Section V presents the results of numerical simulations. Finally, conclusions are drawn in Section VI.

II. Literature Review

It is important to highlight that, in the literature, there exists a very wide variety of methods and studies in the context of blind inversion where the input of the Wiener system is originally assumed to be independent and identically distributed (i.i.d.). Table I gives an overview and comparison of the most recently available blind inversion methods for nonlinear systems. The left column indicates the authors and the inversion criterion, while the right column describes the nature of the input signal, of the Hammerstein inversion model (parametric or non-parametric) and of the optimization procedure.

In spite of the aforementioned efforts, the inversion task for sources with statistical (temporal) dependence still demands further studies. If the samples of $s(n)$ are *dependent* — a quite realistic assumption — the previously mentioned approaches are not capable of obtaining the desired solution, since they consider maximal independence-based criteria to estimate the Hammerstein system optimal parameters. It should be mentioned that dependent sources are practically important in view of, for example, the potential application of different types of codes before signal transmission. Moreover, improvements on problem parametrization and on representation of the candidate solutions are required to overcome stability and convergence issues that may rise when adopting, as inverse filters, a model with feedback loops.

TABLE I: Comparison of recently available blind inversion methods for nonlinear systems.

Model (approach)	Methods
Wiener-Hammerstein (Taleb *et al.*, 2001 [1]). Mutual information	Input: non-Gaussian i.i.d. Invertible nonlinearity and filter Quasi-nonparametric
Wiener-Hammerstein (Solé-Casals *et al.*, 2002 [10]). Mutual information	Input: non-Gaussian i.i.d. Artificial neural networks or polynomials Parametric
Wiener-Hammerstein (Babaie-Zadeh *et al.*, 2003 [11]). Mutual information	Input: non-Gaussian i.i.d. Invertible nonlinearity and filter Minimization projection
Wiener-Hammerstein (Zhang and Chan, 2004 [12] and Sol Casals *et al*. 2005 [13]). Mutual information	Input: non-Gaussian i.i.d. Gaussianization techniques Parametric
Wiener-Hammerstein (Rojas *et al.*, 2007 [14]). Kurtosis	Input: non-Gaussian i.i.d. Polynomial parametric Genetic algorithm
Wiener-Hammerstein (Solé-Casals and Caiafa, 2013 [15]). Mutual information	Input: non-Gaussian i.i.d. Unknown and invertible filter Parametric Accelerated implementation algorithm
Wiener-Hammerstein (Silva *et al.*, 2015 [16]). Mutual information	Input: non-Gaussian i.i.d. Hammerstein system with FIR or IIR structure Immune-inspired optimization algorithms

III. Autocorrentropy Function

Correntropy or, more specifically, the autocorrentropy function was first introduced by Santamaria *et al.* [17], who suggested its application in the blind deconvolution problem. It is a measure that generalizes the autocorrelation function to nonlinear spaces: if $\{X(n), n \in N\}$ is a stochastic process within an index set N, then the autocorrentropy function $V(i_1, i_2)$ is

$$V(i_1, i_2) = E\left[\kappa_\sigma\left(X(i_1) - X(i_2)\right)\right], \tag{3}$$

where $E[\cdot]$ denotes the statistical expectation and $k_\sigma(\cdot)$ is the kernel function, usually assumed to be a Gaussian kernel, given by

$$\kappa_\sigma(x - y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-y)^2}{2\sigma^2}\right), \tag{4}$$

where σ is a parameter known as the kernel size. This parameter plays a crucial role, since it influences in the nature of the performance surface, like the cost function smoothness, the presence of local optima and convergence rate [18]. Using a Taylor series expansion for the Gaussian kernel, it can be shown that correntropy encompasses HOS information and, consequently, can be a more robust dependence measure than autocorrelation function [17].

If $\{X(i)\}$ is stationary, Eq. (3) is simplified to

$$V(m) = E\left[\kappa_\sigma\left(X(i_1+m) - X(i_1)\right)\right], \tag{5}$$

where $m = (i_2 - i_1)$ is the time lag between the samples. It is possible to estimate the autocorrentropy function through the sample mean:

$$\hat{V}(m) = \frac{1}{N-m+1}\sum_{n=0}^{N-m}\kappa_\sigma\left(x(n+m) - x(n)\right). \tag{6}$$

where N is the size of the data window used to estimate the correntropy and $x(n)$ is the data samples available $\{x(n),\ n = 0, 1, \ldots N-1\}$

One can find several applications of correntropy in different domains e.g. nonlinear regression, equalization, blind source separation, independence tests [19] etc. For blind deconvolution of stationary signals, the authors of [17] propose a correntropy-based criterion which minimizes the objective function

$$J_{cor}(\mathbf{w}) = \sum_{m=1}^{P}(V_s(m) - V_y(m))^2, \tag{7}$$

where $\mathbf{w}$ is the filter parameters vector and P is the number of lags. Note that the lag $m = 0$ is not considered, since it is always equal to $k(0) = 1/(\sqrt{2\pi}\sigma)$. In other words, this criterion tries to match the correntropy $V_s(m)$ associated with the source $s(n)$ to the correntropy $V_y(m)$ of the equalizer output $y(n)$.

IV. Proposal of an Immune-Inspired, Correntropy-Based Framework

We consider the Hammerstein system, defined in Eq. (2), to invert the Wiener system. It comprises an invertible nonlinear function $g[\cdot]$, which is assumed to be an odd-power polynomial of $(2N_p - 1)$-th order with strictly positive coefficients

$$g(x) = c_1x^1 + c_2x^3 + \ldots + c_{N_p}x^{2N_p-1}, \quad c_k \geq 0, \quad k = 1, 2, \ldots, N_p \tag{8}$$

followed by an LTI sub-system with infinite impulse response (IIR) $w(n)$. Although some authors adopt finite impulse response (FIR) models for the linear sub-system [1], the choice of the IIR model is justified by its greater flexibility and efficiency in comparison with the FIR structure.

The transfer function of an IIR filter is

$$W(z) = \frac{\sum_{i=0}^{N_a} a_i z^{-i}}{\sum_{k=0}^{N_b} b_k z^{-k}}, \tag{9}$$

where a_k and b_k are the adjustable coefficients of the model. Such parametrization is widely adopted for IIR filter design [16], [20] due to its simplicity and straightforwardness to calculate the filter's frequency response. However, there is a problem: for any choice on values for a_i and b_k, it is necessary to check if the filter is *stable*, i.e. the poles of $W(z)$ are inside the unit-circle. In the context of searching for optimal filter parameters via a population-based metaheuristic, this representation requires a validation (or penalty) routine step for every single candidate-solution and involves an unnecessary search over regions of the parameter space that may result, actually, in non-stable and consequently useless filters.

Hence, this work proposes an alternative model to represent the IIR filter that is based on the zeros and poles of the system [21], [22]:

$$W(z) = a_0\frac{\prod_{i=1}^{2M}(z - r_i)}{\prod_{k=1}^{2M}(z - p_k)} \tag{10}$$

where $N_a = N_b = 2M$ is the order of the filter, $r_1, r_2, \ldots, r_{2M}$ are the zeros, $p_1, p_2, \ldots, p_{2M}$ are the poles and a_0 is the filter gain, which is set to 1 due to the scale indeterminacy of the problem [1]. Each zero and pole is represented by its magnitude and angle, however, only a filter with *real-valued* coefficients is feasible, which is valid when, for any complex-valued zero / pole, its complex conjugate is also a zero / pole. Then it is necessary to include in the encoding vector just M zeros and M poles while their complex conjugates are implicitly considered. A magnitude less than 1 for the poles assures the stability condition, moreover, this restriction is extended for the zeros because we assume that the filter inverse (the linear part of the Wiener system) should be stable, as well.

Fig. 2 summarizes the parameters that each population individual encodes and compares it with the strategy formerly employed in [16], [20], where the real coefficients expressed by Eq. (9) were supposed to be explicitly optimized. Observe that the zeros and poles are represented by the polar notation, where the angle is limited to $[0, \pi]$[1]. The vector dimension is $N_p + 4M$ elements.

To work with this parametric model, in view of the difficulties of the gradient-based methods for nonlinear-multimodal cost functions and, consequently, to escape from local optima, Artificial Immune Systems (AIS) can be an effective search procedure. Due to the successful results presented by CLONALG in [9], [16], this work continues to use this technique as the optimization method, but with the new zero/pole encoding for the IIR filter sub-system, in association with the cost function defined in Eq. (7): the matching of correntropies.

The clonal selection principle was initially based on works carried in the 1970 by Burnett [23]. This work served as inspiration for CLONALG [24], a popular AIS algorithm involving an abstract version of the cloning and hypermutation process. All clonal selection-based algorithms essentially

[1] There is no need to search for negative angles, because of the implicit complex conjugate zero/pole.

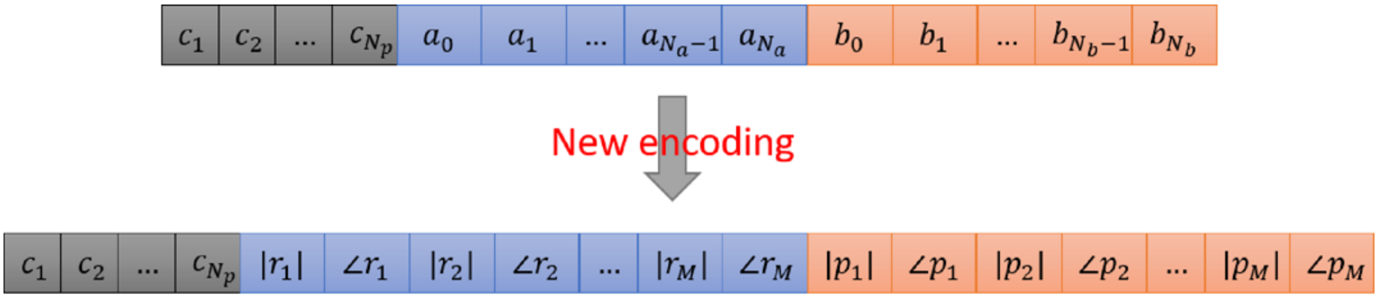

Fig. 2: New proposal of parameter vector of each CLONALG individual, in comparison to the former approach.

Algorithm 1 Pseudo-code of CLONALG algorithm for optimization

Require: [Ab] = clonalg($\beta, \rho, Ninitial, nC, b, range$)
Ensure: Ab = random($Ninitial, range$)
while iteration $\geq maxIT$ **do**
  Solve fit = affinity(Ab)
  C = clone(Ab, nC, β)
  C^* = mutate(C, fit, ρ)
  Fit' = affinity(C^*)
  R = select(C^*, Fit')
  Ab = replace(R, random($b, range$))
end while

gravitate around a repeated cycle of match, clone, mutate and replace, and numerous parameters can be tuned, including the cloning rate, the initial number of antibodies, and the mutation rate for the clones.

The CLONALG algorithm, described in Algorithm 1, is first initialized with an Ab pool of antibodies with fixed size $Ninitial$, in which every Ab_i representing an element from the parameter space. It is proposed by [24] that the generation of those antibodies occurs randomly in order to have a great diversity of population. First, all Ab members are evaluated by the fitness function $f^{Ag}(Ab_i)$, in which Ag represents the antigens, then, the amount of clones nC to be generated for each individual is calculated as follows:

$$nC = \text{round}(\beta \cdot Ninitial), \tag{11}$$

where β is the clonal factor and $Ninitial$ is the size of the antibody pool. Following, the new set of nC clones performs an affinity maturation process. The parameter ρ controls the shape of the mutation rate with respect to the following equation:

$$\alpha = e^{(-\rho \cdot fit)}, \tag{12}$$

where α represents the mutation rate, and fit is the fitness function $f^{Ag}(Ab_i)$ value normalized in $[0, 1]$. Note that the mutation rate is inversely proportional to their parent's affinity, i.e. the greater the affinity, the lower the mutation intensity. Next, a new set of R individuals is formed. In this process, which represents the implementation of the immune memory of the system, the individuals with the highest affinities and diversity are kept. Finally, the main loop is concluded with a random generation of b new antibodies that will replace the lowest affinity Ab in the current population. The process repeats itself until a number of iterations $maxiT$ is executed.

In this paper, the CLONALG optimization algorithm is responsible for searching the optimal parameters of $g[\cdot]$ and $w(n)$ that minimize the cost function J_{cor}, evaluated according Eq. (7). The individuals of the population represent the parameters of the Hammerstein system, according to the parametric models previously defined in Eqs. (8) and (10).

V. Numerical Simulations

For the sake of experimental validation, this section presents a performance evaluation of CLONALG in two sets of experiments, considering continuous and discrete input signals. Moreover, equivalent results employing the Hill Climbing search procedure [25] are exhibited as a baseline, i.e. they work as a reference for testing whether, despite the same cost function, CLONALG's adoption can contribute to obtain better inverse solutions than a simpler search heuristic does.

To reduce indeterminacies, at every fitness evaluation, the output of the nonlinear stage $u(n)$, $n = 1, 2, \ldots, N$ is centered and normalized as well as the output of the linear stage $y(n)$, as stated in [9], [16]. For all scenarios, $N = 2000$ samples of $s(n)$ are considered, where the resulting signal $x(n)$ is provided to the algorithm. The number of lags used in the $J_{cor}(\cdot)$ cost function for equalization is $P = 10$, the kernel size adjustment is discussed in the following section.

CLONALG parameters — the clonal factor (denoted as β), the parameter ρ and the number of individuals $Ninitial$ — were defined similarly to [9], with the aid of a preliminary cross-validation routine comprising 10 independent trials of the algorithm, with the correntropy cost function, for each possible configuration: $\beta \in \{0.1, 0.2, 0.3\}$, $\rho \in \{2, 3, 4, \ldots, 8\}$ and $Ninitial \in \{20, 50, 100\}$. The parameters were fixed at 0.1, 4 and 20 individuals, respectively. The algorithm has 10% of new individuals inserted per iteration and stops when it reaches 100 consecutive generations without improvement on the best solution.

In order to evaluate the method performance, we compute mean values, over a number of independent algorithm executions for each experiment, of the signal to noise ratio (SNR) between the output signal and the original signal for the optimal equalization delay, i.e. $\sigma_s^2/\sigma_n^2 = E[y^2(n)]/E[(s(n) - y(n))^2]$, where σ_n^2 is the error power and σ_s^2 is the estimated signal power.

A. Continuous Case

First, we consider the input signal to be either a Uniform or Laplacian i.i.d. sequence, with zero mean and unit variance, that is submitted to a linear precoder $P(z) = 1 + 1z^{-1}$. The linear precoder can be a type of line coding introducing correlation. Thereby, an i.i.d. source is linearly precoded to form a sequence of dependent samples $s(n)$ [17]. The source autocorrentropy is estimated from a reference set of 500 samples of $s(n)$. The actually transmitted samples of $s(n)$ do not belong to this set. Then, we analyze the performance of the algorithm in a series of scenarios varying the parameters of the Wiener system (the linear $H(z)$ and nonlinear distortion $f[\cdot]$) as well as the order of the Hammerstein linear sub-system $W(z)$.

Before starting the analysis, we empirically search for an appropriate kernel size to be used in the autocorrentropy criterion. We consider 6 possible kernel size values, 0.125, 0.25, 0.5, 1, 2 and 4, and test the performance of the proposed framework considering both source distributions. Two different scenarios for the Wiener system are analyzed:

1) A minimum phase system with coefficients $H(z) = 1 + 0.5z^{-1}$ and nonlinear distortion $f(v) = \text{sign}(v)\sqrt[3]{|v|}$. The polynomial model is set to $N_p = 3$ and the IIR linear sub-system parameter is set to $M = 1$.
2) A higher order linear sub-system $H(z) = 1 - 0.0919z^{-1} + 0.2282z^{-2} - 0.1274z^{-3} + 0.1408z^{-4} - 0.0189z^{-5} + 0.0173z^{-6} - 0.0072z^{-7} + 0.0038z^{-8}$ and a harder nonlinear distortion, $f(v) = (0.1 * v) + \tanh(3v)$; consequently we increment the flexibility of the Hammerstein system model by setting $N_p = 5$ and $M = 2$ i.e. a 4 poles/zeros IIR filter.

The mean results over 50 independent algorithm executions are presented in Fig. 3. Observe that the peak performance is achieved with $\sigma = 4$ for the more complex scenario and uniform-distributed signal. However, in order to define a consensual σ value to use in the following experiments, we pick $\sigma = 2$ because it provides the best *average* SNR value — 13.3312 dB — among the four combinations of distribution and Wiener system setup.

After adjusting the kernel size σ of the autocorrentropy cost function, we proceed to the test of the CLONALG algorithm. Besides the two scenarios that were previously introduced in the kernel bandwidth experiment, two other possibilites are considered:

3) $H(z) = 1 - 0.0919z^{-1} + 0.2282z^{-2} - 0.1274z^{-3} + 0.1408z^{-4} - 0.0189z^{-5} + 0.0173z^{-6} - 0.0072z^{-7} + 0.0038z^{-8}$ and $f(v) = \text{sign}(v)\sqrt[3]{|v|}$. The Hammerstein polynomial model is adjusted to $N_p = 3$ for this case.
4) $H(z) = 1 + 0.5z^{-1}$, and the harder nonlinear distortion, $f(v) = (0.1 * e) + \tanh(3v)$. The Hammerstein polynomial model is adjusted to $N_p = 5$ for this case.

The average results over 50 independent algorithm executions are presented in Table II. As can be seen, both CLONALG and Hill Climbing frameworks are able to compensate the distortions with reasonably performance in terms of the SNR levels. For the source with uniformly distributed sequence, the Hill Climbing metaheuristic outperforms CLONALG for scenarios 1 and 4, which present short-length $H(z)$ – the difference is of about 1 dB and 0.5 dB in each case. However, the CLONALG metaheuristic is able to achieve higher SNR levels when M is required to be larger, i.e. when $H(z)$ presents a larger impulse response, as occurs in scenarios 2 and 3 – now, the difference is of about 1.5 dB and 2.15 dB in each case. For the Laplacian source, the results are similar, however, in average, the performance is slightly reduced. Again, the CLONALG outperforms the Hill Climbing metaheuristic for scenarios 2 and 3, which are more complex. This strongly suggests that the CLONALG framework is more adequate to explore the search space in complex scenarios, while the Hill Climbing method is preferred for the simple ones.

B. Discrete Case

For the discrete case, we consider that the input is (*i*) an i.i.d. signal with samples drawn from the alphabet $\{-1, +1\}$ submitted to the same linear precoder $P(z)$ or (*ii*) the Alternate Mark Inversion (AMI) source [26], whose dependent symbol sequence is drawn from the alphabet $\{-1, 0, +1\}$. The autocorrentropy function of both signals are analytically given by [17]. The number of lags used in the $J_{cor}(\cdot)$ cost function for equalization is $P = 10$.

Similarly to the continuous scenarios, we repeat the validation routine with respect to the kernel size. Fig. 4 indicates that the algorithm obtained the best average result with the kernel size $\sigma = 0.125$ for the four scenarios — 27.4944 dB —, hence this will be our choice for the following experiments

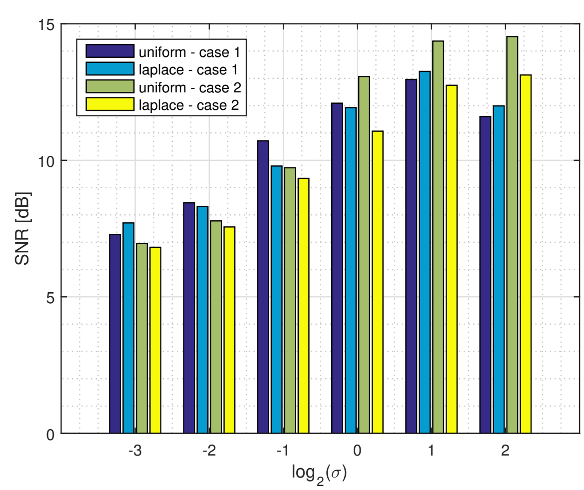

Fig. 3: Kernel performance for the continuous source case, inverting two different scenarios of Wiener system.

TABLE II: SNR Performance results for the continuous scenario. Top and bottom values of each cell correspond to the CLONALG and Hill Climbing framework, respectively.

Uniform		M=1	M=2	M=3	M=4
Case 1	CLONALG	12.8824	12.5903	12.7683	12.4268
	HC	**13.8159**	12.8625	12.1602	9.6512
Case 2	CLONALG	14.0067	**14.1585**	13.8982	13.7879
	HC	12.2639	12.3202	12.5708	11.3178
Case 3	CLONALG	13.1099	**14.0176**	13.9986	13.5933
	HC	11.3457	11.8466	11.3059	11.3594
Case 4	CLONALG	13.5644	13.2425	13.0783	13.1814
	HC	**14.0935**	13.8996	12.5805	11.4232

Laplace		M=1	M=2	M=3	M=4
Case 1	CLONALG	13.8492	12.6269	12.6800	12.2896
	HC	**14.3684**	11.5424	10.0331	7.7197
Case 2	CLONALG	12.8849	**12.9363**	12.6426	12.5975
	HC	12.0423	11.8657	9.8045	9.4489
Case 3	CLONALG	13.4433	13.4973	13.1930	**13.5461**
	HC	12.8281	11.6275	10.0896	8.8354
Case 4	CLONALG	11.0794	10.4045	9.5882	9.4390
	HC	**11.4265**	9.7858	7.9453	7.5232

with discrete-valued signals. In comparison to the continuous-valued signals, we can point out that the algorithm showed a better performance with smaller kernel sizes, an explanation for this behavior may lie on the fact that for those particular signals, while the observed values are spread on a continuous scale due to the infinite feedback of the IIR, the expected values should cluster strongly in the vicinity of the original alphabet values and, thus, the correntropy corresponds to the coincidence count of those discrete values in lagged and paired replicates of the signal, within a small neighboring coincidence region as 0.1. This idea reinforces the interpretation of correntropy as a coincidence detection problem [27].

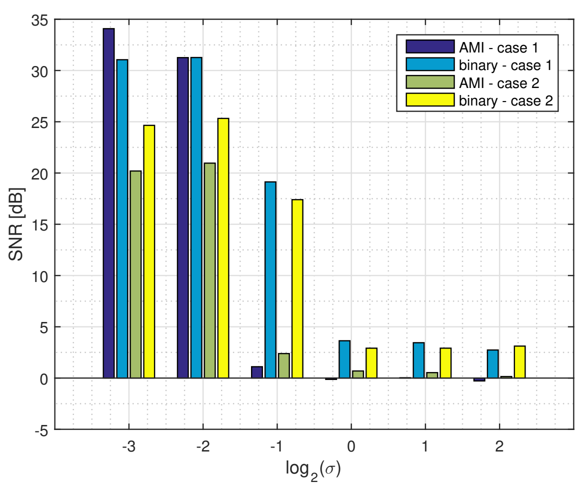

Fig. 4: Kernel performance for the discrete source case, inverting two different scenarios of Wiener system.

Now, the same scenarios employed in the continuous case are considered to evaluate CLONALG algorithm. Recall that $\sigma = 0.125$. The results are shown in Table III.

TABLE III: SNR Performance results for the discrete scenario. Top and bottom values of each cell correspond to the CLONALG and Hill Climbing framework, respectively.

AMI		M=1	M=2	M=3	M=4
Case 1	CLONALG	33.9124	**36.2919**	35.3603	31.8558
	HC	11.7922	9.8823	9.1711	6.0548
Case 2	CLONALG	**23.1519**	20.6934	18.7298	10.0731
	HC	5.0663	1.8992	1.4138	-0.7963
Case 3	CLONALG	**24.4768**	23.0702	15.9888	10.8499
	HC	5.6252	0.6495	-0.8191	-1.7857
Case 4	CLONALG	30.8552	**32.3474**	30.8517	26.5868
	HC	9.8691	9.4982	9.2538	6.0676

Binary		M=1	M=2	M=3	M=4
Case 1	CLONALG	32.1971	**33.8023**	31.5411	22.7820
	HC	15.7949	10.1553	4.9099	3.0888
Case 2	CLONALG	22.4573	24.2831	**24.4593**	20.8216
	HC	19.2722	19.4053	11.8491	8.1095
Case 3	CLONALG	17.1621	**18.4213**	11.4710	3.3564
	HC	14.3148	8.6529	2.0810	0.1145
Case 4	CLONALG	32.0605	**32.3550**	29.5918	21.5619
	HC	13.4862	7.7855	1.9758	1.4910

It is possible to note that the achieved SNR level are higher than that obtained in the continuous case. For the AMI case, the Hill Climbing method has found certain difficulty in exploring the search space, performing poorer in all four scenarios considered in comparison with the CLONALG – the higher SNR value for the Hill Climbing is 11.8 dB, while for the CLONALG is 36.3 dB. The CLONALG metaheuristic, on the other hand, was able to find good solutions with $M = 1$ even for the complex scenarios 2 and 3. For binary source, the Hill Climbing metaheuristic improves its performance in comparison with the previous source type, but still performs poorer than the CLONALG framework. Indeed, the CLONALG metaheuristic is able to find good quality solutions for $M \geq 2$, while the Hill Climbing solutions are better for $M = 1$. These results corroborate the idea the CLONALG metaheuristic is capable of exploring complex search spaces with more efficiency than the Hill Climbing.

In a general perspective, the correntropy-based criterion provided good results that validate the idea of using statistical dependence as criterion to invert the original system. Also, one can see that the CLONALG-based algorithm presented better results than Hill Climbing: for all cases in the discrete scenario and for most of the cases in the continuous scenario. Furthermore, it is possible to see that the feedback loop in the linear filter was pertinent to build up the inversion performance, enhancing the flexibility on modeling the Hammerstein system.

VI. Final Remarks

In this work, we proposed the use of a framework composed of the immune-inspired metaheuristic CLONALG and of the correntropy-based criterion for the unsupervised Wiener-Hammerstein problem. Particularly, a new antibody/cell representation based on the zeros and poles of the linear stage of the Hammerstein system was considered, in order to avoid solutions with stability issues.

A series of simulations were carried out in four different scenarios to analyze the performance of the proposed framework, as well as the kernel size sensitivity. The results indicated that a larger kernel size (with $\sigma = 2$) is preferred for continuous sources, while a smaller kernel size ($\sigma = 0.125$) is more suitable for discrete sources. The proposed framework was able to outperform the baseline search method (Hill Climbing) in several cases, mainly when the scenario is more complex and requires a more efficient exploration of larger search spaces. In addition, the CLONALG metaheuristic also exhibited a considerably better performance for discrete sources, confirming the adequacy of CLONALG immune-inspired algorithm as optimization strategy for this task.

As future work, it is clear that additional analysis of the framework are still necessary in terms of other signals, distortion models and noisy scenarios, while CLONALG efficiency has to be compared with more recent AIS optimization techniques.

ACKNOWLEDGMENT

The authors would like to thank CNPq, FAPESP (2017/11488-5) and CAPES for the financial support.

REFERENCES

[1] A. Taleb, J. Solé-Casals, and C. Jutten, "Quasi-nonparametric Blind Inversion of Wiener Systems," *IEEE Transactions on Signal Processing*, vol. 49, no. 5, pp. 917–924, 2001.

[2] O. Goryachkin and A. Berezovskiy, "Blind signal processing in telecommunication systems based on polynomial statistics," *American Journal of Computational Mathematics*, vol. 4, no. 03, pp. 233–241, 2014.

[3] V. Zarzoso, "Blind and semi-blind signal processing for telecommunications and biomedical engineering," Ph.D. dissertation, Université Nice Sophia Antipolis, 2009.

[4] R. Bars, I. Bèzi, B. Pilipar, and B. Ojhelyi, "Nonlinear and Long Range Control of a Distillation Pilot Plant," *Identification and System Parameter Estimation. 9 th IFAC/IFORS Symposium*, pp. 848–853, 1990.

[5] R. C. Brinker, "A Comparison of Results from Parameter Estimations of Impulse Responses of the Transient Visual System," *Biological Cybernetics*, vol. 61, pp. 139–151, 1989.

[6] J. Príncipe, D. Xu, and J. Fisher, "Information Theoretic Learning," *Unsupervised adaptive filtering*, vol. 1, pp. 265–319, 2000.

[7] D. Erdogmus and J. Príncipe, "An Error-Entropy Minimization Algorithm for Supervised Training of Nonlinear Adaptive Systems," *IEEE Transactions on Signal Processing*, vol. 50, no. 7, pp. 1780–1786, 2002.

[8] K. E. Hild, D. Erdogmus, and J. Príncipe, "Blind Source Separation Using Renyi's Mutual Information," *IEEE Signal Processing Letters*, vol. 8, no. 6, pp. 174–176, 2001.

[9] S. A. Fernandez, R. Attux, D. Fantinato, J. Montalvão, and D. G. Silva, "An immune-inspired, dependence-based approach to blind inversion of wiener systems," in *ESANN 2016 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning.*, 2016, pp. 201–206.

[10] J. Solé-Casals, C. Jutten, and A. Taleb, "Parametric Approach to Blind Deconvolution of Nonlinear Channels," *Neurocomputing*, vol. 48, no. 1, pp. 339–355, 2002.

[11] M. Babaie-Zadeh, J. Solé-Casals, and C. Jutten, "Blind Inversion of Wiener System Using a Minimization-Projection (MP) Approach," *ICA [S.l.: s.n.]*, pp. 681–686, 2003.

[12] L. Chan, "Practical Method for Blind Inversion of Wiener Systems," in *IEEE International Joint Conference on Neural Networks*, vol. 3, 2004, pp. 2163–2168.

[13] J. Solé-Casals, C. Jutten, and D. Pham, "Fast Approximation of Nonlinearities for Improving Inversion Algorithms of PNL Mixtures and Wiener Systems," *Transactions on Signal processing*, vol. 85, no. 9, pp. 1780–1786, 2005.

[14] F. Rojas, J. Solé-Casals, and C. G. Puntonet, "An Evolutionary Approach for Blind Inversion of Wiener Systems," in *Independent Component Analysis and Signal Separation*. Springer, 2007, pp. 260–267.

[15] J. Solé-Casals and C. Caiafa, "A Fast Gradient Approximation for Nonlinear Blind Signal Processing," *Cognitive Computation*, vol. 5, no. 4, pp. 483–492, 2013.

[16] D. Silva, J. Montalvão, R. Attux, and L. Coradine, "An immune-Inspired, Information-Theoretic Framework for Blind Inversion of Wiener Systems," *Signal Processing*, vol. 113, pp. 18–31, 2015.

[17] I. Santamaría, P. Pokharel, and J. Príncipe, "Generalized Correlation Function: Definition, Properties, and Application to Blind Equalization," *IEEE Transactions on Signal Processing*, vol. 54, no. 6, pp. 2187–2197, 2006.

[18] S. Abhishek and J. Príncipe, "Information Theoretic Learning with Adaptive Kernels," *Signal Processing*, vol. 91, no. 2, pp. 203 – 213, 2011.

[19] J. Príncipe, *Information Theoretic Learning: Renyi's Entropy and Kernel Perspectives*. Springer Science & Business Media, 2010.

[20] Anonymous, "Work related to the present work," in *ESANN 2016 proceedings*, 2016, pp. 201–206.

[21] R. Storn, "Differential evolution design of an iir-filter," in *Evolutionary Computation, 1996., Proceedings of IEEE International Conference on*. IEEE, 1996, pp. 268–273.

[22] D. R. de Oliveira and H. S. Lopes, "Codificação por pólos e zeros para o projeto de filtros digitais de resposta infinita ao impulso utilizando otimização por enxame de partículas," in *XIX Congresso Brasileiro de Automática - CBA 2012*, Campina Grande, 2012.

[23] F. M. Burnet, "Clonal Selection and After," *Theoritical Immunology*, pp. 63–85, 1978.

[24] L. N. de Castro and F. J. Von Zuben, "Learning and Optimization Using the Clonal Selection Principle," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 3, pp. 239–251, Jun 2002.

[25] P. Langley, J. Gennari, and W. Iba, "Hill climbing theories of learning," in *Proceedings of the Fourth International Workshop on Machine Learning*, 1987, pp. 312–323.

[26] J. G. Proakis, *Intersymbol Interference in Digital Communication Systems*. Wiley Online Library, 2001.

[27] J. Montalvão, J. Canuto, and E. Carvalho, "A correntropy function based on coincidence detection," *Pattern Recognition Letters*, vol. 85, pp. 84–88, 2017.