

# An Immune-Inspired, Dependence-Based Approach to Blind Inversion of Wiener Systems

Stephanie Milena Alvarez Fernandez<sup>1</sup>, Romis Attux<sup>2</sup>,  
Denis G. Fantinato<sup>2</sup>, Jugurta Montalvão<sup>3</sup> and Daniel G. Silva<sup>1</sup>

1- University of Brasilia - Dept. of Electrical Engineering

2- University of Campinas - School of Electrical and Computer Engineering

3- Federal University of Sergipe - Dept. of Electrical Engineering

**Abstract.** In this work, we present a comparative analysis of two methods — based on the autocorrelation and autocorentropy functions — for representing the time structure of a given signal in the context of the unsupervised inversion of Wiener systems by Hammerstein systems. Linear stages with and without feedback are considered and an immune-inspired algorithm is used to allow parameter optimization without the need for manipulating the cost function, and also with a significant probability of global convergence. The results indicate that both functions provide effective means for system inversion and also illustrate the effect of linear feedback on the overall system performance.

## 1 Introduction

Adaptive inverse modeling is an important task in view of its vast horizon of practical applications in the signal processing area. This is due to the necessity of understanding, analyzing, predicting and controlling real systems, which has grown quickly with the technological and industrial advance [1]. Particularly in this work, we consider the inverse modeling of an important model, the Wiener system, which consists of a linear time-invariant (LTI) filter subsystem  $h(n)$  followed by a memoryless, invertible, nonlinear distortion  $f[\cdot]$ :

$$x(n) = f[e(n)] = f[h(n) * s(n)], \quad (1)$$

where  $s(n)$  is the system input signal and  $x(n)$  is its output system. Despite its simplicity, it can be applied within many contexts [1, 2], by providing a simple mathematical treatment for handling nonlinear dynamics. Various methods have been developed for the task of Wiener system inversion, employing concepts derived from linear optimization, nonparametric regression, and nonlinear optimization with models such as polynomials, neural networks and orthogonal functions [3]. Particularly, one straightforward structure to invert a Wiener system is the well-known Hammerstein system, which is composed of a static nonlinear block  $g[\cdot]$  followed by a LTI subsystem  $w(n)$ :

$$y(n) = w(n) * u(n) = w(n) * g[x(n)], \quad (2)$$

where  $x(n)$  is the input signal to the Hammerstein system and  $y(n)$  its output. Taleb et al. [4] and Silva et al. [5] proposes blind methods to invert Wiener systems through Hammerstein models. Consider that  $h(n)$ ,  $f[\cdot]$  and  $s(n)$  are unknown, but the latter is assumed to be composed of independent and identically distributed (i.i.d.) samples: in this context, the inversion problem consists

of finding  $w(n)$  and  $g[\cdot]$  such that the output  $y(n)$  of the Hammerstein system be as close as possible to the original signal  $s(n)$ , despite a scale and time delay ambiguity. Hence, this task is implemented via a maximal independence criteria to estimate the optimal parameters.

However, if the samples of  $s(n)$  are dependent, the previously mentioned approaches are not capable of inverting the system. In this case, it is plausible to adopt some dependence measure as criterion for the inversion task. Additionally, by assuming statistically dependent samples, the scope of acceptable input signals is expanded to include, for instance, encoded signals.

In the context of Information Theoretic Learning (ITL) [6], a new generalized correlation function, called correntropy, has been introduced. Along with the autocorrelation function, they share the property of taking into account the time structure of a random process, but, differently from autocorrelation, correntropy is not limited to second order moments. With this in mind, this work proposes a modified version of the original method introduced in [5], for blind inversion of Wiener systems, by considering the adoption of autocorrelation or correntropy as dependence-based criteria, in association with an immune-inspired search algorithm.

The rest of the paper is divided as follows: Section 2 details the two dependence criteria to be comparatively studied through the remainder of the work; Section 3 presents a brief introduction to the immune-inspired algorithm and its application; Section 4 presents the results of numerical simulations and, finally, conclusions are drawn in Section 5.

## 2 Dependence Measures

Correntropy or, more specifically, the autocorrentropy function was first introduced by Santamaria et al. [7], who suggested an initial application to blind deconvolution. It is a measure that generalizes the autocorrelation function to nonlinear spaces: if  $\{x_n, n \in \mathcal{N}\}$  is a discrete-time, stationary stochastic process within an index set  $\mathcal{N}$ , then the autocorrentropy function is defined as

$$V_x(m) = E[k_\sigma(x_n - x_{n-m})], \quad (3)$$

where  $E[\cdot]$  denotes the statistical expectation and  $k_\sigma(\cdot)$  is generally the Gaussian kernel function [8], with  $\sigma$  being the parameter known as the kernel size.

Using a series expansion for the Gaussian kernel, it can be shown that correntropy contains higher order information and, consequently, can be a more robust dependence measure between time samples in comparison with the autocorrelation function [7]. Given that the probability description of a process is commonly unknown, correntropy can be easily estimated via the sample mean over  $N$  samples of  $\{x_n\}$ :

$$\hat{V}_x(m) = \frac{1}{N - m + 1} \sum_{n=m}^N k_\sigma(x_n - x_{n-m}). \quad (4)$$

One can find several applications of correntropy in different domains, e.g. nonlinear regression, equalization and blind source separation [8]. For blind

deconvolution, the authors of [7] proposed a correntropy-based criterion which comprises minimizing the objective function:

$$J_{cor}(\theta) = \sum_{m=1}^P (V_s(m) - V_y(m))^2, \quad (5)$$

where  $\theta$  represent the deconvolution filter parameters and  $P$  is the number of lags. This criterion tries to match the correntropy  $V_s(m)$  associated with the source  $s(n)$  to the correntropy  $V_y(m)$  of the filter output  $y(n)$ .

The “classical” dependence measure, the autocorrelation function, is easily defined for  $\{x_n\}$  as  $R_x(m) = E[x_n x_{n-m}]$ . Analogously, an autocorrelation-based criterion [9] is straightforward to define:

$$J_R(\theta) = \sum_{m=1}^P (R_s(m) - R_y(m))^2. \quad (6)$$

In the context of this work, where  $s(n)$  is an input to the Wiener system and  $y(n)$  is the output of the Hammerstein system, we propose, analogously to the linear blind deconvolution problem, to employ Eqs. (5) and (6) as criteria to estimate the inverse model. The idea is that the temporal dependence signature of the original signal provides sufficient information to estimate the Hammerstein system and, consequently, to obtain an estimate for  $s(n)$ .

### 3 Algorithm

As already mentioned, we consider the Hammerstein system defined in (2), to invert the Wiener system. The structural components of the Hammerstein system are comprised of a nonlinear function  $g[\cdot]$ , which is assumed to be an odd-power polynomial of  $(2k - 1)$ -th order with strictly positive coefficients  $g(x) = c_1 x^1 + c_2 x^3 + \dots + c_k x^{2k-1}$ ,  $c_k \geq 0, \forall k$ , followed by an LTI sub-system with impulse response  $w(n)$ , which is represented by a finite number of coefficients that describes its transfer function. Despite the adoption by some authors of a finite impulse response (FIR) model for the linear sub-system [10], we employ, in consonance with [5], a more powerful linear structure — the infinite impulse response (IIR) filter —, with input-output relationship

$$y(n) = \sum_{k=0}^M a_k u(n-k) + \sum_{k=1}^Q b_k y(n-k), \quad (7)$$

where  $M + 1$  and  $Q$  are the number of adjustable coefficients  $a_k$  and  $b_k$  of the model. Note that an immune-inspired algorithm is an appropriate search procedure to work with this parametric model, due to the difficulties of gradient-based methods to (i) avoid stability issues with IIR filters and (ii) achieving a local optimum.

The identification of Hammerstein systems via Artificial Immune Systems — specifically, the CLONALG algorithm [11] — was already considered in [5]. Due to the successful results presented by the technique in this previous proposal, this work maintains CLONALG as the optimization method, but of a different criteria, as mentioned in Section 2: the matching of correntropies and autocorrelations (see Eqs. (5) and (6)).

---

**Algorithm 1** Pseudo-code of CLONALG algorithm for optimization

---

**Require:**  $[Ab] = \text{clonalg}(N_{\text{initial}}, nC, b, \text{range})$   
**Ensure:**  $Ab = \text{random}(N_{\text{initial}}, \text{range})$   
1: **while** iteration  $\geq \text{maxIT}$  **do**  
2:     Solve  $\text{fit} = \text{affinity}(Ab)$   
3:      $C = \text{clone}(Ab, nC)$   
4:      $C^* = \text{mutate}(C, \text{fit})$   
5:      $\text{Fit}' = \text{affinity}(C^*)$   
6:      $P = \text{select}(C^*, \text{Fit}')$   
7:      $Ab = \text{replace}(P, \text{random}(b, \text{range}))$   
8: **end while**

---

The CLONALG algorithm is inspired in the Clonal Selection principle, and is characterized by a population of antibodies,  $Ab$ , whose affinity (or fitness) with respect to the antigen  $Ag$  is represented by the objective function. The algorithm is initialized with an  $Ab$  pool of fixed size  $N_{\text{initial}}$ , in which every  $Ab_i$  represents an element from the parameter space, possibly an optimal solution. First, the fitness function evaluates  $f^{Ag}(Ab_i)$  for all  $Ab$  members, then, it proceeds by selecting a subset of  $n$  antibodies that have the highest affinities, which are subsequently cloned. The set of  $nC$  clones then experience an affinity maturation process, where the intensity of modifications is inversely proportional to their parent's affinity. In the sequence, the clones are compared to their parent in order to select the one with the highest affinity. 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  $\text{maxIT}$  is executed. After that, the solution is the best individual of  $Ab$  (higher affinity). Algorithm 1 presents the pseudo-code of CLONALG; for more details, we recommend [5, 11].

To summarize, the CLONALG algorithm with real encoding is responsible for searching the optimal parameters of  $g[\cdot]$  and  $w(n)$  that minimize the cost function,  $J_{\text{cor}}$  or  $J_R$ , evaluated according Eqs. (5) and (6). The individuals of the population represent the parameters of the Hammerstein system, according to the parametric models previously defined in the beginning of this section.

## 4 Numerical Simulations

This section tests the proposal performance in two scenarios, with input signals comprising dependent samples, continuously- or discretely-distributed. First, we consider  $s(n)$  to be an uniform i.i.d. sequence submitted to a linear precoder with transfer function  $P(z) = 1 + 1z^{-1}$  [7]. Its autocorrentropy and autocorrelation are estimated from 500 samples, which are distinct from the (unknown to the algorithm) input to be submitted to the Wiener system. For the discrete case,  $s(n)$  is an i.i.d. signal with samples drawn from the alphabet  $\{-1, +1\}$  submitted to the same linear precoder  $P(z)$ , whose autocorrentropy functions is analitically given in [7] and  $R_s(m) = \sigma_s^2 \delta(m-1) + 2\sigma_s^2 \delta(m) + \sigma_s^2 \delta(m+1)$ , where  $\sigma_s^2$  is its variance. Finally, we consider the Alternate Mark Inversion (AMI) source, whose dependent symbol sequence is drawn from the alphabet  $\{-1, 0, +1\}$ , its autocorrentropy function is also defined in [7] and  $R_s(m) = -\frac{1}{2}\sigma_s^2 \delta(m-1) +$

$\sigma_s^2 \delta(m) - \frac{1}{2} \sigma_s^2 \delta(m+1)$ . For all scenarios,  $N = 2000$  unit-variance samples of  $s(n)$  are considered, where the resulting signal  $x(n)$  is provided to the algorithm. The number of lags is  $P = 10$  and the kernel size is defined as  $\sigma = 0.4$ .

The algorithm is tested considering distinct  $M$  and  $Q$  values. The CLONALG parameters were adjusted to the values suggested in [5], i.e. 300 iterations, 50 individuals, and 10% of new individuals inserted per iteration. Other CLONALG parameters were defined with the aid of a preliminary cross-validation routine, which comprised 10 independent trials of the algorithm, with the correntropy cost function, for each possible configuration: the clone number parameter  $\beta \in \{0.1, 0.2, 0.3\}$  and the mutation rate  $\rho \in \{2, 3, 4, \dots, 8\}$  [11]. The selected values were  $\beta = 0.1$  and  $\rho = 4$ .

The performance is measured by the output signal to noise ratio  $\text{SNR} = 10 \log_{10} E[y^2(n)]/E[(s(n) - y(n))^2]$  averaged over a set of 10 independent trials of the algorithm. In a first scenario, the Wiener system is modeled by a minimum phase system with coefficients  $H(z) = 1 + 0.5z^{-1}$  and nonlinear distortion  $f(e) = \text{sign}(e) \sqrt[3]{|e|}$ , the polynomial model is set to  $k = 3$ . The second scenario analyzes the algorithm in a more complex situation,  $f(e) = \tanh(3e)$  and  $H(z) = 1 - 0.0919z^{-1} + 0.2282z^{-2} - 0.1274z^{-3} + 0.1408z^{-4} - 0.0189z^{-5} + 0.0173z^{-6} - 0.0072z^{-7} + 0.0038z^{-8}$ , with the polynomial model set to  $k = 5$ . Table 1 shows the mean SNR for the three distributions in both scenarios. A highlighted value corresponds to the best result for each criterion.

In a general perspective, the autocorrelation and the correntropy-based criteria provided good performances in inverting the original system. One can see that the correntropy-based criterion presented the best results: for all cases in the first scenarios and for the precoded binary signal in both scenarios; however, this criterion showed an inferior overall performance for the continuous signal and AMI line code in the 2nd scenario, where autocorrelation performed better. Furthermore, it is possible to see that the feedback loop in the linear filter was pertinent to build up the inversion performance, since most of the top scores, for both criteria, were obtained with  $Q \geq 1$ . In terms of criteria, it is not conclusive which one of the two possibilities is preferable, hence demanding a more thorough analysis. Withal, both indicate that the exploration of the temporal structure of the input signals can lead to the inversion of Wiener systems.

## 5 Final Remarks

This work presented a comparison between two strategies for quantifying the time structure of a signal in the context of the problem of unsupervised Wiener-Hammerstein inversion: the autocorrelation and autocorrentropy functions. The possibility of linear feedback was taken into account and different simulation scenarios were considered. An immune-inspired algorithm was employed in view of the characteristics of the associated optimization problem. The results have shown that both statistical methods are feasible, and also revealed the potential of using feedback loops.

		1st Scenario			2nd Scenario		
		M			Q		
		0	1	2	0	1	2
Uniform	1	N/A	<b>14.8568</b> <b>14.6424</b>	10.5581 12.1894	N/A	6.6408 <b>8.4698</b>	<b>7.1136</b> 7.5135
	2	6.9127 9.6851	6.5298 9.9486	6.5864 8.0404	5.3037 5.6166	3.5200 4.6532	3.4118 4.3149
	3	4.2240 7.1719	7.8640 7.2597	7.0788 6.1361	5.9823 4.1090	2.7248 3.1194	2.7801 3.1099
	1	N/A	<b>42.3971</b> 8.3664	34.1152 6.0335	N/A	27.1582 <b>11.4609</b>	263274 6.8779
	2	15.4638 9.3447	15.0029 7.2518	33.4532 8.7425	27.2379 5.6971	27.0057 6.1671	<b>28.3819</b> 7.9212
	3	24.5195 <b>10.1324</b>	26.0103 6.6126	27.8254 7.8646	22.3939 6.3156	25.6450 6.4826	26.1593 8.7940
AMI	1	N/A	<b>76.5697</b> 14.8651	36.7313 <b>16.5883</b>	N/A	<b>5.6105</b> <b>8.7712</b>	1.4308 8.6935
	2	5.2648 6.5509	15.6042 14.7809	11.5516 10.7013	0.7853 4.1404	0.1777 4.2678	3.6273 5.3170
	3	18.4655 7.9706	5.6490 8.4627	11.8754 10.2360	3.4488 3.1774	4.1253 4.0420	3.0500 5.3144

Table 1: Results for the first and second scenarios. Top values of each cell correspond to the technique with  $J_{cor}(\cdot)$  and the lower values with  $J_R(\cdot)$ .

## References

- [1] B. Widrow and E. Walach. *Adaptive Inverse Control, Reissue Edition: A Signal Processing Approach*. John Wiley & Sons, 2008.
- [2] I. W. Hunter. The identification of nonlinear biological systems: Wiener and hammerstein cascade models. *Biological cybernetics*, 55:135–144, 1986.
- [3] T. Nitta. Orthogonality of decision boundaries in complex-valued neural networks. *Neural computation*, 16(1):73–97, 2004.
- [4] A. Taleb, J. Solé-Casals, C. Jutten, et al. *Quasi-nonparametric blind inversion of Wiener systems*. Institute of Electrical and Electronic Engineers, IEEE, 2001.
- [5] D. G. Silva, J. Montalvão, R. Attux, and L. Coradine. An immune-inspired, information-theoretic framework for blind inversion of wiener systems. *Signal Processing*, 113:18–31, 2015.
- [6] D. Erdogmus and J. Príncipe. An error-entropy minimization algorithm for supervised training of nonlinear adaptive systems. *Signal Processing, IEEE Transactions on*, 50(7):1780–1786, 2002.
- [7] I. Santamaría, P. Pokharel, and J. Príncipe. Generalized correlation function: definition, properties, and application to blind equalization. *Signal Processing, IEEE Transactions on*, 54(6):2187–2197, 2006.
- [8] J. Príncipe. *Information theoretic learning: Renyi’s entropy and kernel perspectives*. Springer Science & Business Media, 2010.
- [9] D. Fantinato, R. R. Attux, A. Neves, R. Suyama, and J. Romano. Blind Deconvolution of Correlated Sources Based on Second-Order Statistics. In *SBrT 2013*, pages 1–5. Sociedade Brasileira de Telecomunicações, 2013.
- [10] A. Taleb, J. Solé-Casals, and C. Jutten. Quasi-nonparametric blind inversion of wiener systems. *Signal Processing, IEEE Transactions on*, 49(5):917–924, May 2001.
- [11] L. N. de Castro and F. J. Von Zuben. Learning and optimization using the clonal selection principle. *Evolutionary Computation, IEEE Transactions on*, 6(3):239–251, Jun 2002.