Non-Observability Analysis by Means of an Evolutionary Technique in Measured Electric Power Systems

Santiago Vazquez-Rodriguez*†‡, Jesús Á. Gomollón†, Gervasio Varela*†, Alejandro Paz-Lopez*†
*Integrated Group for Engineering Research
†University of A Coruña
c/Mendizábal s/n, 15403 - Ferrol, Spain
‡Email: svr@udc.es

***Abstract*—The observability characterization of an Electric Power System (EPS) from a topological point of view with respect to a given measurement acquisition system is equivalent to the existence of a certain spanning tree. In previous work, a genetic algorithm was developed in order to address this issue. In this paper the behavior of this evolutionary algorithm is studied by means of probabilistic methods. Although the main purpose of the algorithm is to find a tree, the determination of the non-existence of the tree due to the uncertainty inherent to evolutionary techniques was addressed using statistical hypothesis testing. This allows to characterize, with a given certainty level, the non-existence of such a solution and, therefore, the non-observability of the EPS. The techniques developed in this paper were tested over two benchmark systems: the IEEE networks with 118 and 300 nodes.**

I. Introduction

The places over a region or a country where electric energy is generated (sources) and where it is consumed (loads or drains) are separated hundreds or thousands of kilometers from each other. *Electric Power Systems* (EPS) are in charge of guaranteing that electricity is available wherever it is demanded. The most visible elements in an EPS are the high voltage transportation lines and the substations where these lines are incident. The resulting topology is typically a mesh and is known as *electric power network*. Therefore, graphical terminology is commonly used to describe the network, where sources, loads and power transformers are identified with the nodes while branches are the lines that join those elements.

Such a system is constrained by the equations derived from electrical circuit theory, the network topology and the value of the parameters that characterize the system. The *state variables* are defined as a minimum set of electrical variables from which any other system variable can be calculated by means of the equation system. Thus, it is said that the state of a system is known when the entire state variables are also known, and the process to evaluate it is known as *state estimation*, first described by Schweppe et al. in [1].

In order to address the problem of estimating the state of a system, it is necessary to acquire a collection of electric measurements throughout the network that allow formulating a sufficient number of independent equations. A system variable is said to be *observed* if it is directly measured in the network. A system variable is said to be *observable* with respect to a set of measurements if it can be estimated from the available data. By extension, an EPS is said to be *observable* with respect to a given measurement acquisition system if the entire state of the system can be estimated. Otherwise, the system is said to be *not observable* or *unobservable*.

Two main strategies can be distinguished in the literature in order to address the observability analysis in EPS. On one hand, numerical and algebraic methods, starting with the contribution of Monticelli and Wu in [2], where the approximate model suggested in [1] is assumed. In this model, only node voltages and active and reactive power injections and flows are considered in the measurement acquisition system, grouped in two categories: *node measurements* when they are associated to nodes and *branch measurements* when they are acquired in network branches. Besides the techniques proposed in [2] other algorithms were developed [3][4], numerical observability analysis was extended into new considerations [5][6] and other variables where included in the measurement set [7][8].

On the other hand, topological approaches arise from the work of Krumpholz et al. in [9], where the authors establish a necessary and sufficient condition for *topological observability* by means of the existence of certain *spanning tree* within the system network. This tree must be constructed from the measurement set considered in such a way that each branch is associated to one measured variable in accordance to a series of assignment rules. Thus, any technique focused on topological observability determination must search for any spanning tree among all graphs included in the system network that fit the assignment rules depending on a given measurement set. Different approaches to this issue can be found in the literature: combinatorial [10], combinatorial and matroids to characterize graphs [11], neural networks [12][13] and evolutionary techniques [14][15][16].

In previous work [16], an evolutionary technique was developed for which the behavior of the algorithm was studied in terms of convergence, that is, how good the algorithm was at finding a valid spanning tree and how fast it was. In spite of the fact that the convergence was reached in a high percentage of cases, a natural question arises: what happens when the algorithm does not find any valid spanning tree? May it be concluded that the system is not observable? The response to this last question is, of course not, mainly because all the tests were run over observable configurations.

The aim of the present work is to study the non-observability condition and how it can be characterized by means of evolutionary techniques in conjunction with statistical theory. For that purpose, the IEEE networks with 118 and 300 nodes were considered as benchmark systems for which up to one hundred measurement configurations were defined. In order to characterize the behavior of the algorithm statistically, 150,000 simulations were run over the above systems and measurement configurations. In this paper, the results of these tests are briefly presented and some conclusions are reached about the characterization of non-observability.

The rest of the paper is organized as follows. In section II a summary of the main points dealt with in [16] is shown. In section III the benchmark networks and their measurement configurations considered in tests are introduced. Section IV is devoted to describing how the tests were designed and programed and how and what data should be taken into account. In section V a statistical hypothesis testing is introduced in order to characterize the non-observability of a system. A summary of the results reached in the work is shown in section VI. Finally, conclusions are presented in section VII.

II. Previous Work

All the measurements taken into account in [16] and in the present work are categorized as branch measurements or node measurements. The first group consists of active and reactive electric powers that flow through a branch and, therefore, whose values are measured at that branch. Node voltage measurements and active and reactive powers injected in a node are grouped as node measurements. Figure 1 shows an example in which a seven node network represents the topology of an electric power system. A total of six power measurements were numbered and highlighted with thicker lines in the figure: three node measurements, ordered from 1 to 3 at nodes 2, 4 and 6, respectively; and three branch measurements, ordered from 4 to 6. It is also assumed that the voltage of at least one of the network nodes is known.

A. Graph Construction

Given an electric power network and a measurement acquisition system, a set of measurement assignment rules are described in [9] in order to construct graphs in the network, as follows:

- Each branch measurement is assigned to the branch of the network corresponding to its position.
- Each node measurement is assigned to one and only one of the branches incident to the node.

Any graph constructed in such a way is known as a *graph of full rank* and the search space for the evolutionary algorithm is made up of all the graphs of full rank within the network.

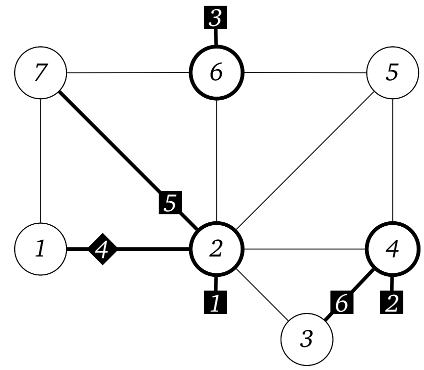


Fig. 1. A seven node network example and a collection of three node and three branch measurements

Figure 2 shows a graph of full rank belonging to the search space resulting from the above mentioned example and its associated measurement system. Each edge in the graph is assigned to one of the measurements considered. Note that oriented edges are stated in the figure in order to clarify the measurement from which those branches are derived.

B. Encoding

Each gene in the adopted encoding scheme consists of an integer value that points to the branch assigned to a given node measurement. Thus, a chromosome has as many genes as node measurements. Note that the branches associated to branch measurements are contained in all the graphs in the search space.

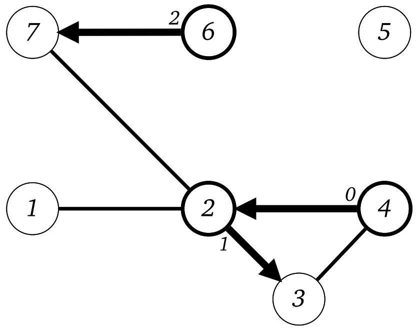
The numbers displayed in Figure 2 close to the edges derived from node measurements correspond to an arbitrary numbering of branches previously held in the range from zero up to the maximum number of node incident branches minus one. These collection of integer values form the graph encoding, as shown in the figure.

C. Fitness Criteria

A fitness vector made up of three integer values has been defined insted of a fitness function. If two graphs must be compared, the first indices of the fitness vector are checked. Only when these are equal are the second ones compared, and so on. These three indices have been defined as follows:

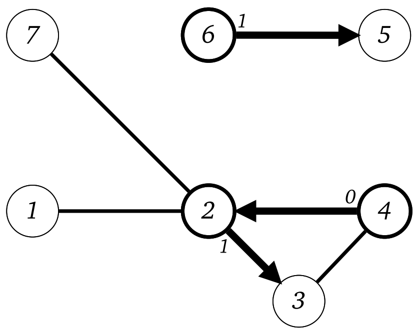
1) ind_1 = *nodes in the graph* − *subgraphs in the graph*
2) ind_2 = *number of subgraphs in the graph*
3) ind_3 = *nodes of the smallest subgraph in the graph*

Since the graph in Figure 2 consists of six connected nodes in a unique subgraph, the values of the aforementioned fitness indices are given by 5, 1 and 6, respectively. The variation of a gene in the encoding is equivalent to the reassignment of the associated node measurements to a different network branch, as shown in Figure 3, which results in a different graph and, occasionally, a different fitness vector. In particular the graph in Figure 3 is considered a better qualified graph than the previous one because the isolated node 5 has joined another node, in spite of the fact that the largest subgraph is smaller



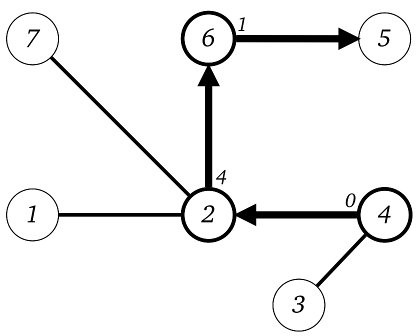
$$chromosome = [\ 1 \quad 0 \quad 2\ ]$$
$$fitness\ vector = [\ 5 \quad 1 \quad 6\ ]$$

Fig. 2. Graph of full rank with six connected nodes



$$chromosome = [\ 1 \quad 0 \quad 1\ ]$$
$$fitness\ vector = [\ 5 \quad 2 \quad 2\ ]$$

Fig. 3. Graph of full rank with two subgraphs



$$chromosome = [\ 4 \quad 0 \quad 1\ ]$$
$$fitness\ vector = [\ 6 \quad 1 \quad 7\ ]$$

Fig. 4. Spanning tree of full rank

than in Figure 2. Eventually, the evolutionary algorithm may reach a graph like the one in Figure 4, which joins all the network nodes in an unique subgraph.

III. Test Cases

The IEEE 118 and IEEE 300 node networks used in these tests are frequently used as benchmark EPS in literature. These data can be found in [17]. In order to characterize the behavior of the evolutionary algorithm, it was necessary to design a large enough number of measurement configurations and, therefore, reproduce a statistically significant case set in tests. The idea consists in characterizing the non-observability condition by means of the characterization of the behavior over observable cases, by exclusion. This is the reason why all the measurement configurations considered are observable. Therefore, the evolutionary algorithm is said to have converged when any spanning tree of full rank is reached in the search space. In what follows, the number of generations needed for convergence will be the subject of study. Convergences will be taken into account if they take place in a fixed number of generations which is denoted as *sample window*. Table I shows a summary of the networks used in the tests as benchmark cases.

TABLE I
Benchmark Networks

Network	*Nodes*	*Branches*	*Measure config*	*Sample window*	*Carried out simulations*
IEEE-118	118	179	50	50	100,000
IEEE-300	300	409	50	70	50,000

The measurement configurations were deliberately designed to subject the algorithm to the most complex scenarios and, in particular, much more complex than in real cases. Hence, the results are more pessimistic than in real environments. That is true because of the following two reasons:

- The measurement configurations considered are, all of them, *critically observable*, that is, there are no *redundant* measurements.All the measures are said to be *critical* because the lack of any one would make the system unobservable. Due to the fact that a spanning tree has as many branches as nodes minus one and any branch is associated to one measurement, the number of measurements in each configuration is equal to the number of network nodes minus one.
- Most of the measurements are node measurements instead of branch ones, and the larger the number of node measurements, the larger the size of the chromosome. Figure 5 shows in logarithmic scale the size of the search space for each of the measurement configurations for the IEEE 118 node network. The number of node measurements grows with the order of the configuration system, resulting in spaces in the range of 9.4×10^{25} up to 3.6×10^{50} graphs of full rank. The larger ones corresponding to a configuration with 117 node measurements. In the case of the IEEE 300 node network, the number of graphs in the search spaces goes from 2.2×10^{57} up to 2.6×10^{106}.

In general, the measurement acquisition systems in real EPS are clearly redundant and there exist a significant number of branch measurements.

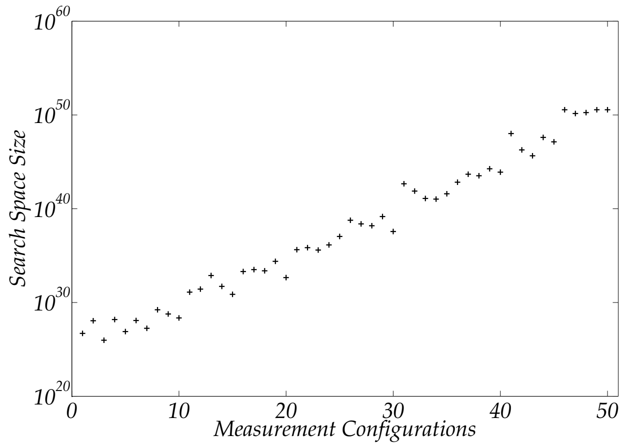


Fig. 5. Search Space Size Resulting for the Evolutionary Algorithm and the 50 Measurement Configurations Designed for the IEEE 118 Node Network

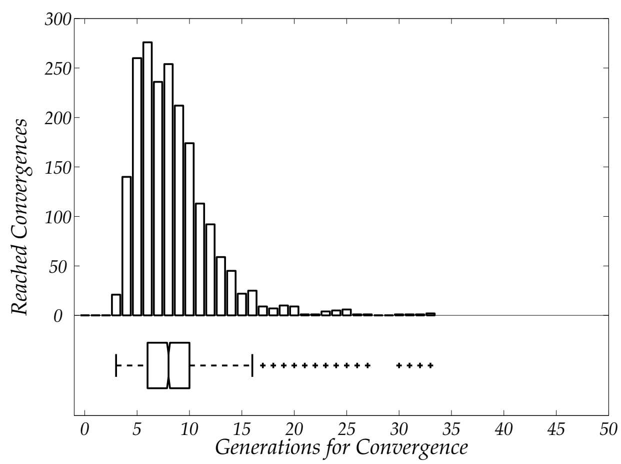


Fig. 6. Algorithm Frequency Convergence for the IEEE 118 Node Network and a Given Measurement Configuration with 77 Node Measurements and 40 Branch Measurements

IV. Programmed Tests

Starting with the IEEE 118 node network, a pilot sampling was carried out to determine how the sample space seems to be. For that purpose, the Chebyshev's inequality was considered:

$$P\left(|X-\mu|<k\sigma\right)\geq 1-\frac{1}{k^2} \tag{1}$$

where X denotes the random variable, that is, the number of generations needed to reach the convergence of the algorithm, μ is the finite expected value of X, σ denotes the non-zero standard deviation and k is a positive real number. After these pilot results, it could be seen that in 75% of cases, the convergence of the algorithm had occurred before the first 30 generations and the expected number of generations needed was around 10 generations. Therefore, from (1) it can be established that the standard deviation σ should be less than 10 generations.

Thus, a sample window of 50 generations was fixed and a minimum sample size n can be determined by minimizing the

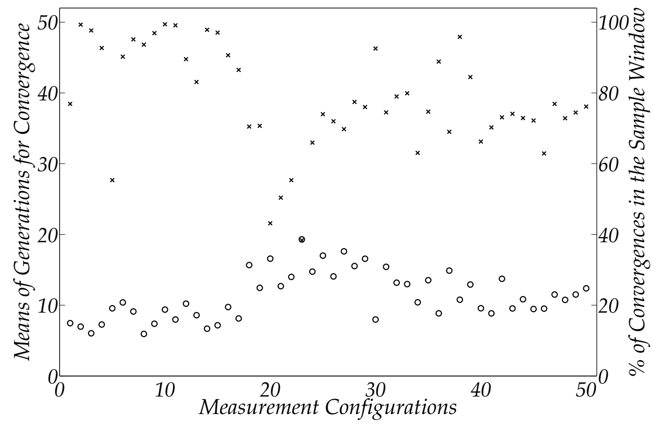


Fig. 7. Convergence of the Evolutionary Algorithm after Running 50 *Measurement Configurations* $\times$ 2000 *Simulations* over the IEEE 118 Node Network

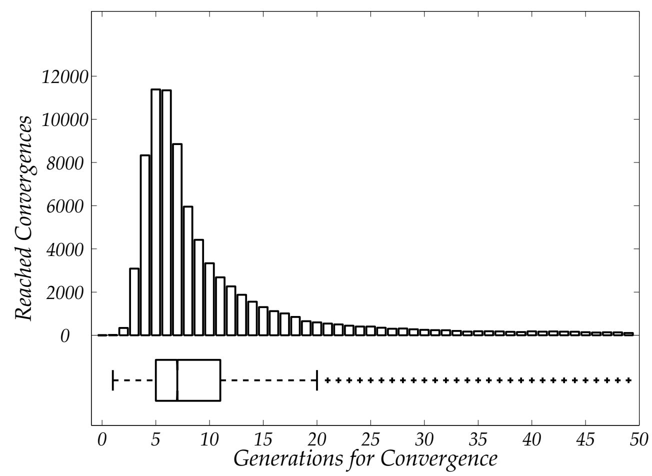


Fig. 8. Algorithm Frequency Convergence after Running 50 *Measurement Configurations* $\times$ 2000 *Simulations* over the IEEE 118 Node Network

error in the estimation of the mean of X for a given *tolerable error* $\alpha=0.05$ from the equation:

$$\epsilon=\pm z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}} \tag{2}$$

resulting in an error of $\epsilon=\pm 0.44$ generations for a sample size of $n=2000$ simulations.

One of the main contributions emphasized in [16] was the good behavior of the evolutionary algorithm in terms of the number of generations needed to reach convergence. Figure 6 shows the frequency convergence diagram after evolving populations of 1000 graphs a total number of 2000 times for a given measurement configuration over the IEEE 118 node network. A box plot of the generations needed for convergence is also stated in the figure in order to point out the median and the 25th and 75th percentiles.

In brief, the evolutionary algorithm was run 2000 times for each of the 50 measurement configurations designed for the IEEE 118 node network. The entire number of simulations carried out was 100,000, in which populations of 1000 graphs

TABLE II
SUMMARY OF RESULTS OBTAINED IN THE TESTS WITH IEEE 118 AND 300 NODE NETWORKS

Network	$\bar{S}$	$\bar{X}$	$\epsilon_{\bar{X}}$	$\bar{p}$	$\epsilon_{\bar{p}}$	n_p	$(1-p)^{n_p}$
IEEE-118	7.41	11.26	± 2.05	0.75	± 0.12	4	0.0041
IEEE-300	15.57	30.88	± 4.32	0.45	± 0.14	8	0.0087

TABLE III
STATISTICAL HYPOTHESIS TESTING, TRUE PANNEL

Null Hypothesis	*The System is Observable*	*The System is Not Observable*
Accepted	Right	Type II Error
Rejected	Type I Error	Right

were evolved. The convergence was registered when it was reached in the sample window of the first 50 generations. Figure 7 shows convergence results for each of the 50 measurment configurations (horizontal axis, ordered from 1 to 50), where circles denote the mean number of generations estimated for the convergence in the given sample window (left scale) and crosses denote the percentage of convergences reached in the sample window (right scale). Figure 8 shows how the behavior of the algorithm when all the measurement configurations are taken into account is the same as the one shown in Figure 6 for a given measurement system.

Note that the values shown in Figure 7 are, in themselves, a new sampling of algorithm convergence for any configuration of measurements, and from this new data the estimation of the generations expected to reach convergence $\bar{X}$ and the error of this estimation $\epsilon_{\bar{X}}$ can be calculated. Let $\bar{x}_i$ and $\bar{s}_i$ be the estimated mean and standard deviation, respectively, after having run the entire simulations with the i-th measurement configuration. Then, after $n_m = 50$ batteries of tests, one for each measurement system, it follows that

$$\bar{X} = \frac{1}{n_m} \sum_{i=1}^{n_m} \bar{x}_i \tag{3}$$

and the sampling standard deviation $\bar{S}$ is given by

$$\bar{S} = \frac{1}{n_m} \sum_{i=1}^{n_m} \bar{s}_i \tag{4}$$

that allows to calculate $\epsilon_{\bar{X}}$ from (2). The values obtained for these statistical variables for a given tolerable error $\alpha = 0.05$ are shown in Table II.

V. NON-OBSERVABILITY CHARACTERIZATION

In this section, the non-observability characterization will be achieved by means of statistical hypothesis testing. For this purpose, the probability of the algorithm converging in a sample window when the system is observable is first analyzed.

Let's consider the IEEE 118 node network and let p_i be the probability of reaching the convergence of the algorithm in the sample window for the i-th measurement configuration. The estimation of the probability of convergence for any observable configuration $\bar{p}$ can be given by the harmonic mean as follows:

$$\bar{p} = n_m \left(\sum_{i=1}^{n_m} \frac{1}{p_i} \right)^{-1} \tag{5}$$

and the error of this estimation is given by:

$$\epsilon_{\bar{p}} = \pm z_{\alpha/2} \sqrt{\frac{\bar{p}\,(1-\bar{p})}{n_m}} \tag{6}$$

From the values of these variables shown in Table II for the IEEE 118 node network and a given tolerable error $\alpha = 0.05$, it follows that, for any observable measurement configuration, there exists 95% percent probability that the per unit number of cases to reach the convergence in the first 50 generations is limited by:

$$L = (0.63; 0.87) \tag{7}$$

where L is the confidence interval of $\bar{p}$ for a given tolerable error α.

Note that this result allows characterizing in statistical terms the non-observability of an EPS by means of hypothesis testing.

Let a measurement configuration be considered for which the observability condition is not known. The evolutionary algorithm is run a given number of times n_p in such a way that the processes can be considered as independent events. Let p be the per unit number of cases that the algorithm converges in the sample window. A *null hypothesis* can be established:

$$H_0 = \textit{The system is observable} \tag{8}$$

that must be contrasted with the *alternative hypothesis* given by

$$H_1 = \textit{The system is not observable} \tag{9}$$

Then, depending on accepting or rejecting the null hypothesis, one of four situations can take place, as shown in Table III. The scenarios where the diagnosis is wrong are of interest in terms of probabilities of that occurring. Let the next two cases be considered:

- *The system is not observable*: In this case, no algorithm will reach convergence and, then, the probability of concluding that the system is observable is null. In other words, the probability of a *type II error* being committed is zero.
- *The system is observable*: The probability of a given process reaching convergence in the sample window is $1-p$, while the probability of no process converging is reduced to $(1-p)^{n_p}$. As a result, the significance level of the test, that is, the probability of incorrectly rejecting the null hypothesis, which is known as *type I error* is $(1-p)^{n_p}$.

At the same time and independently of the probability of convergence taking place, the registration of a convergence in any of the processes characterizes unambiguously and with certainty the observability of the system.

In summary, after the statistical analysis of the behavior of the evolutionary algorithm over a given EPS, it is possible to characterize non-observability conditions of any measurement configuration by means of a number of independent processes n_p and statistical hypothesis testing with a significance level limited by $(1-p)^{n_p}$.

Table II shows that 4 independent process are necessary to reach a significance level of the test of less than 1% with the IEEE 118 node network. In other words, the statistical study carried out for this network has determined that at least one among four independent processes will reach convergence in the sample window if the measurement configuration is observable. This is true for a given tolerable error but, as the tested measurement configurations were designed much more restrictively than those in real operation scenarios, it is plausible to conclude that the significance level of the hypothesis testing is less than previously mentioned.

Similar tests were carried out with the IEEE 300 node network, for which 50 measurement configurations were designed and where 1000 simulations were run for each measurement system. The populations were of 6000 graphs and they were evolved in a sample window of 70 generations. Table II shows a summary of the results for this network. It can be seen that 8 independent processes bring the significance level of the non-observability characterization to values of less than 1%.

VI. RESULTS

The results achieved in this work can be summarized as follows:

- The evolutionary algorithm has been subjected to multiple measurement scenarios and different EPS and, in all of them, the convergence in terms of number of generations needed to reach the convergence in a sample window was studied. This has allowed to describe the behavior of the algorithm by means of statistical patterns.
- This statistical characterization has been carried out for any given EPS and a large enough number of observable measurement configurations.
- Although the evolutionary algorithm was designed to find a spanning tree of full rank in a search space of graphs, the characterization of the non-existence of such a tree was carried out by means of statistical hypothesis testing. In other words, in spite of the fact that the behavior of any evolutionary algorithm is affected by uncertainty, implementation scenarios were defined in order to characterize, with a given certainty level, the non-existence of a solution.
- Although these techniques were designed for the particular case of measured EPS, they could be put into practice with other real cases in which the non-existence of a solution is relevant.

VII. CONCLUSION

In this paper, a genetic algorithm in conjunction with statistical methods have been put into practice in order to address certain aspects of a particular problem that is not well solved only using evolutionary techniques. In particular, these methods were implemented to address the non-observability analysis of a given EPS with respect to a measurement acquisition system. A series of observable measurement configurations were designed for which batteries of tests were run. The results obtained have allowed to describe the behavior of the evolutionary algorithm for the EPS and any given measurement configuration in terms of statistical patterns. Finally, the non-observability condition characterization was achieved by means of statistical hypothesis testing. All the tests and results shown in this paper were carried out over the IEEE benchmark networks with 118 and 300 nodes. It is plausible to conclude that the techniques developed in this paper are not exclusive to the non-observability of EPS issue and they could be implemented in other real cases.

ACKNOWLEDGMENT

This work was partially funded by the Spanish MICINN through project TIN2011-28753-C02-01 and the Xunta de Galicia and European Regional Development Funds through projects 09DPI012166PR and 10DPI005CT. The technical support of CESGA is gratefully acknowledged.

REFERENCES

[1] F. Schweppe and al., "Power system static-state estimation, part i, ii and iii," *Power Apparatus and Systems IEEE Transactions on*, vol. PAS-89, no. 1, pp. 120–135, 1970.

[2] A. Monticelli and F. Wu, "Network observability: Theory," *Power Apparatus and Systems, IEEE Transactions on*, vol. PAS-104, no. 5, pp. 1042 –1048, may 1985.

[3] A. Monticelli and F. F. Wu, "Observability analysis for orthogonal transformation based state estimation," *Power Systems, IEEE Transactions on*, vol. 1, no. 1, pp. 201 –206, feb. 1986.

[4] F. Wu, W.-H. Liu, L. Holten, L. Gjelsvik, and S. Aam, "Observability analysis and bad data processing for state estimation using hachtel's augmented matrix method," *Power Systems, IEEE Transactions on*, vol. 3, no. 2, pp. 604 –611, may 1988.

[5] P. Katsikas and G. Korres, "Unified observability analysis and measurement placement in generalized state estimation," *Power Systems, IEEE Transactions on*, vol. 18, no. 1, pp. 324 – 333, feb 2003.

[6] J. London, J.B.A., L. Mili, and N. Bretas, "An observability analysis method for a combined parameter and state estimation of a power system," in *Probabilistic Methods Applied to Power Systems, 2004 International Conference on*, sept. 2004, pp. 594 –599.

[7] A. Exposito and A. Abur, "Generalized observability analysis and measurement classification," *Power Systems, IEEE Transactions on*, vol. 13, no. 3, pp. 1090 –1095, aug 1998.

[8] B. Xu and A. Abur, "Observability analysis and measurement placement for systems with pmus," in *Power Systems Conference and Exposition, 2004. IEEE PES*, oct. 2004, pp. 943 – 946 vol.2.

[9] G. Krumpholz, K. Clements, and P. Davis, "Power system observability: A practical algorithm using network topology," *Power Apparatus and Systems, IEEE Transactions on*, vol. PAS-99, no. 4, pp. 1534 –1542, july 1980.

[10] R. Nucera and M. Gilles, "Observability analysis: a new topological algorithm," *Power Systems, IEEE Transactions on*, vol. 6, no. 2, pp. 466 –475, may 1991.

[11] V. Quintana, A. Simoes-Costa, and A. Mandel, "Power system topological observability using a direct graph-theoretic approach," *Power Apparatus and Systems, IEEE Transactions on*, vol. PAS-101, no. 3, pp. 617 –626, march 1982.

[12] H. Mori, “Application of a revised boltzmann machine to topological observability analysis,” in *Neural Networks to Power Systems, 1991., Proceedings of the First International Forum on Applications of*, jul 1991, pp. 283 –287.

[13] A. Jain, J. Choi, and J. Min, “Power system network observability determination using feedforward neural networks,” in *Power System Technology, 2002. Proceedings. PowerCon 2002. International Conference on*, vol. 4, 2002, pp. 2086 – 2090 vol.4.

[14] H. Mori and S. Tsuzuki, “A fast method for topological observability analysis using a minimum spanning tree technique,” *Power Systems, IEEE Transactions on*, vol. 6, no. 2, pp. 491 –500, may 1991.

[15] H. Mori, “A ga-based method for optimizing topological observability index in electric power networks,” in *Evolutionary Computation, 1994. IEEE World Congress on Computational Intelligence., Proceedings of the First IEEE Conference on*, jun 1994, pp. 565 –568 vol.2.

[16] S. Vazquez-Rodriguez, A. Faina, and B. Neira-Duenas, “An evolutionary technique with fast convergence for power system topological observability analysis,” in *Evolutionary Computation, 2006. CEC 2006. IEEE Congress on*, 0-0 2006, pp. 3086 –3090.

[17] (2012, Mar) Power systems test case archive. University of Washington, Dept of Electrical Engineering. [Online]. Available: http://www.ee.washington.edu/research/pstca/index.html