

# Emergent white noise behavior in location independent representations

Ivan I. Garibay<sup>1,2</sup> and Annie S. Wu<sup>1</sup>

<sup>1</sup> University of Central Florida, School of Computer Science,  
P.O. Box 162362, Orlando, FL 32816-2362, USA,  
{igaribay,aswu}@cs.ucf.edu,

WWW home page: <http://ivan.research.ucf.edu>

<sup>2</sup> University of Central Florida, Office of Research  
Orlando Tech Center/ Research Park  
12443 Research Parkway Orlando, FL 32826, USA.

**Abstract.** We use spectral density measurement techniques to analyze the best individuals obtained by a location independent Genetic Algorithm (GA) when applied to three problem domains: number matching, symbolic regression, and dynamical system control. We find that for all the cases the emergent ordering of genomic symbols resembles white noise.

## 1 Introduction

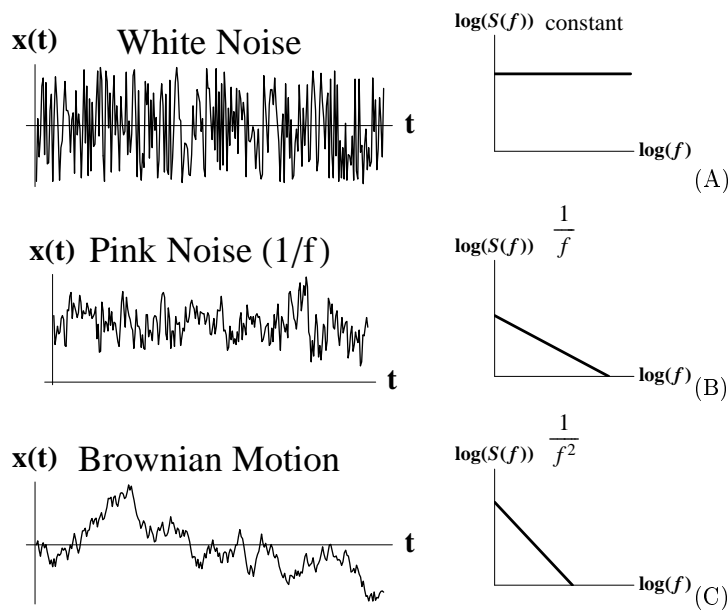
It has been shown using standard spectral density measurement techniques that individual base positions in DNA sequences exhibit  $1/f$  noise and long-range fractal correlations [1]. From the evolutionary computation perspective, we can post the following question: What kind of behavior do genomic symbols present in successfully evolved individuals? After a brief analysis, it is easy to realize that for most evolutionary computation representations, which rely on an order-based encoding of information, this question has a trivial answer: the behavior exhibit by the evolved symbolic sequences is the one dictated by the chosen fitness function. For location independent representations this question becomes more interesting and less obvious. In location independent representations, the fitness function do not take into account the order of the symbols [2]. Therefore, the order of genomic symbols is free to evolve along with the candidate solution they encode. We use spectral density measurement techniques to analyze the best individuals obtained by a location independent Genetic Algorithm (GA)[3, 4, 2, 5] when applied to three problem domains: number matching, symbolic regression, and dynamical system's control. We find that for all the problem domains the emergent ordering of genomic symbols resembles white noise.

## 2 Proportional Representation

In order to study the behavior of location independent representation we use a GA with *proportional representation* [2] for all the experiments. In this representation, encoded information depends solely on what does or does not exist on

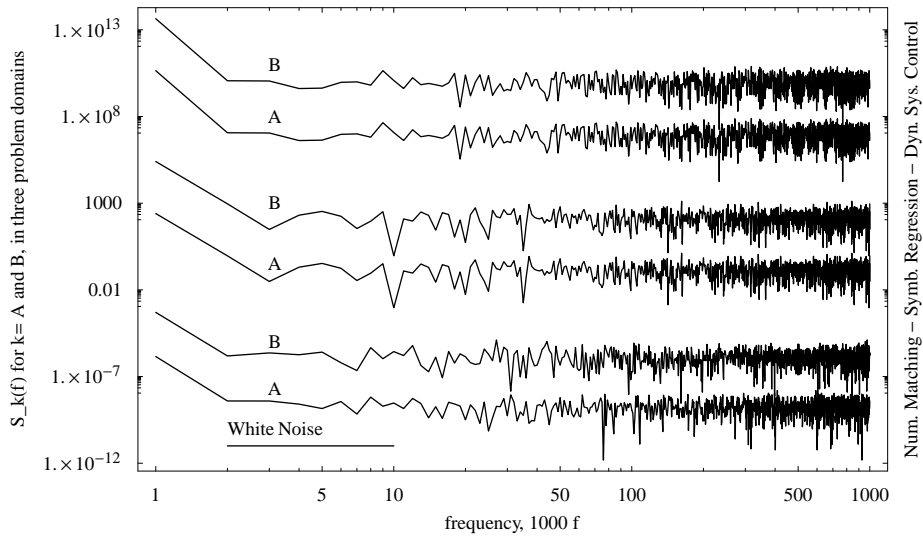
an individual and not on the order on which it is present. As a result, the order of the encoding is free to evolve along with the encoded solution. The proportional representation uses a linear genome with a multicharacter alphabet. One or more unique characters are assigned to each component of a solution or gene. The value of the gene is determined by the relative proportions of the characters assigned to that gene. For full description of the proportional representation and its applications, we refer the reader to [2, 5].

### 3 Symbolic Sequences Analysis



**Fig. 1.** White(A), Pink(B) and Brownian Motion(C) noises and their spectral densities,  $S(f)$ .

*Autocorrelation* and *spectral density* functions are widely used to analyze how fluctuations of a quantity,  $X(t)$ , are correlated between times  $t$  and  $t + \tau$ . Figure 1 shows examples of typical noises and their characteristic spectral densities  $S(f)$ . Voss [1] have shown that this time correlations analysis can be successfully adapted to the analysis of symbolic sequences such DNA. We use this same technique to analyze GA genomes using proportional representation. We analyze the genomes of best of last generation individuals. This individuals are all optimal or are very close to the optimal solution. Figure 2 shows the spectral density of the first two symbols of the alphabet (A and B) for three problems:



**Fig. 2.** Log-log spectral density plots of best of last generation individuals. From bottom up, symbols A and B for PGA on number matching, on symbolic regression and on dynamical system's control problems. The graphics have been off set for clarity.

number matching, symbolic regression, and control of dynamical systems. For all cases, we find that the emergent ordering of genomic symbols, according to our symbolic sequence analysis, resembles white noise.

## Acknowledgment

We thank Richard Voss for useful correspondence.

## References

1. Voss, R.F.:  $1/f$  noise and fractals in DNA-base sequences. In Crilly, Earnshaw, Jones, eds.: *Applications of Fractals and Chaos*. Springer-Verlag (1993) 7–20
2. Wu, A.S., Garibay, I.: The proportional genetic algorithm: Gene expression in a genetic algorithm. *Genetic Programming and Evolvable Hardware* **3** (2002) 157–192
3. Holland, J.H.: *Adaptation in Natural and Artificial Systems*. University of Michigan Press (1975)
4. Goldberg, D.E.: *Genetic algorithms in search, optimization, and machine learning*. Addison Wesley (1989)
5. Wu, A.S., Garibay, I.: Intelligent automated control of life support systems using proportional representations. *IEEE Transactions on Systems, Man, and Cybernetics—Part B* **34** (2004) 1423–1434