Spurious Correlation as an Approximation of the Mutual Information between Redundant Outputs and an Unknown Input

Shuhei Ikemoto[a,*], Fabio DallaLibera[b,1], Koh Hosoda[a], Hiroshi Ishiguro[b]

[a] *Department of Multimedia Engineering, Osaka University,2-1 Yamadaoka, Suita, Osaka, Japan, +81-6-6879-7739*
[b] *Department of Systems Innovation, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka, Japan*

Abstract

Stochastic resonance (SR) is a counterintuitive phenomenon, observed in a wide variety of nonlinear systems, for which the addition of noise of opportune magnitude can improve signal detection. Tuning the noise for maximizing the SR effect is important both for artificial and biological systems. In the case of artificial systems, full exploitation of the SR effect opens the possibility of measuring otherwise unmeasurable signals. In biology, identification of possible SR maximization mechanisms is of great interest for explaining the low-energy high-sensitivity perception capabilities often observed in animals. SR maximization approaches presented in literature use knowledge on the input signal (or stimulus, in the case of living beings), and maximize the mutual information between the input and the output signal. The input signal, however, is unknown in many practical settings. To cope with this problem, this paper introduces an approximation of the input-output mutual information based on the spurious correlation among a set of redundant units. A proof of the approximation, as well as numerical examples of its application are given.

*Corresponding author

Email addresses: `ikemoto@ist.osaka-u.ac.jp` (Shuhei Ikemoto), `fabio.dl@irl.sys.es.osaka-u.ac.jp` (Fabio DallaLibera), `koh.hosoda@ist.osaka-u.ac.jp` (Koh Hosoda), `ishiguro@sys.es.osaka-u.ac.jp` (Hiroshi Ishiguro)

[1] Research Fellow of the Japan Society for the Promotion of Science

Preprint submitted to Communications in Nonlinear Science and Numerical Simulation March 28, 2014

Keywords: stochastic resonance, total correlation

1. Introduction

Stochastic resonance [1] is a counterintentuitive phenomenon, initially proposed in physics for explaining the recurrence of ice ages [2], for which noise takes an important role in enhancing, instead of degrading, the information transmission in nonlinear systems. This phenomenon has been observed in a variety of natural systems [3, 4] and was exploited in a wide range of artificial systems [5, 6, 7, 8, 9, 10, 11].

Among the possible ways of measuring the stochastic resonance effect, mutual information has been proposed as a very theoretically sound approach [12]. In literature, the computation of the mutual information between the conditioning input and the conditioned outputs is often performed assuming the knowledge of the input itself [13, 14] or of its Fisher information [15, 16, 17, 18]. In particular, the computation of the input-output mutual information based on input's Fisher information is an approximation valid in a particular setting, the case in which a single input signal conditions a high number of (redundant) outputs. This setting takes a very important role in engineering applications [8, 7, 10, 19], where a single unknown signal is often measured by multiple sensing elements. The same setting assumes great importance in biology as well [16, 20], because information is often coded and transmitted using populations of neurons [21, 22, 23]. In fact, it was shown that information transmission through multiple low capacity pathways turns out to be more efficient, in terms of metabolic cost, than a single high bit-rate communication channel. The setup of a single input conditioning multiple output is often taken as a first approximation of the real network, that includes coupling between the outputs as well. In this paper we focus on this setup of a single input conditioning multiple outputs, and introduce a new approximation of the input-output mutual information based only on the spurious correlation between a high number of redundant outputs.

The earliest discussion on spurious correlation can be found in [24, 25]. Here, Pearson introduced the term to indicate the correlation that arises in the ratios X/Y and Z/Y even when X, Y, Z are independent random variables. From this, he concludes that correlation between ratios in bone lengths does not constitute a good indication of correct grouping of bones. Later, Yule recognized that the problem does not lie in taking ratios, but that the validity of computing correlations on ratios actually depends on what the variable is expected to influence. In fact, in some cases, it may be possible to expect influence on the absolute magnitudes. In other cases, an effect on the ratios could appear more reasonable. Finally, in some settings the mode of operating of the causes may be totally unknown [26]. In his following works, Yule elaborates on the problem of inferring direct causal relations from associations when no causal relation actually exists. In particular, the specific case of inferring that A causes B when actually A and B are both caused by C, the setup used in this paper, is discussed in his text [27] under the name of "illusory associations". We can find an analysis of the same case, under the currently used name of "spurious correlation", in [28], where Simon clarifies the logical processes and assumptions that are involved in testing whether a correlation is spurious or not. To the best of our knowledge, however, not much research was conducted in studying what kind of information the statistical dependence between several variables influenced by a common variable (or set of variables) can provide on the relationship between the influencing variable(s) and the influenced variables. Nonetheless, it is clear that a link between the two quantities could be exploited in many cases, among which we find the stochastic resonance tuning.

To study this aspect, in this paper we leverage on the concept of total correlation. This quantity was first introduced by McGill in [29]. This work extends Shannon's concept of transmitted information to the case of multiple sources, providing a mathematical model of psychological experiments in which several stimuli influence the measured quantity. Later, the concept of total correlation, also called multiinformation or multivariate constraint, was focused more in detail by several researchers, including Watanabe [30], Garner [31] and Han [32]. In particular, Han gives a full characterization of all the nonnegative symmetric information-theoretic correlation measures, of which total correlation is one instance. This characterization provides a more concrete idea of the measure by showing how total correlation can be decomposed into level-specific measures of dependence among variables. This concept was recently highlighted again by Studeny in [33], where the links between total correlation and conditional mutual information are focused. For the comprehension of the following of the paper, it may suffice to consider total correlation as one of the extensions of mutual information to multiple variables, which measures the deviation of set of variables from independence. More concretely, the total correlation of a set of variables $Y_1, \ldots, Y_n$ can be calculated as the Kullback-Leibler divergence between their joint probability $p(Y_1, \ldots, Y_n)$ and the independent distribution $\prod_i p(Y_i)$. In case of discrete probability distributions we can write

$$C(Y_1, \ldots, Y_n) = \sum_{y_1 \in \mathcal{Y}_1} \cdots \sum_{y_n \in \mathcal{Y}_n} p(y_1, \ldots, y_n) \log \frac{p(y_1, \ldots, y_n)}{p(y_1)p(y_2) \ldots p(y_n)}.$$

In the following, it will be shown that in the setup of a single variable X that influences a set of n variables Y_i, the total correlation between the Y_i variables provides quantitative information on the degree of statistical dependence between X and each of the variables Y_i. This relationship is particularly interesting in all the cases in which X is a latent (unobservable) variable while the values assumed by the variables Y_i can be measured.

In particular, we will prove that the average mutual information between X and each variable Y_i is asymptotically equal to the total correlation among the variables Y_i divided by n when $n \to \infty$. The application of this result will then be discussed in the stochastic resonance setup. It will be shown that this relationship between mutual information and total correlation allows tuning the level of noise without knowing the conditioning signal. The paper concludes with a final remark on the implications of this asymptotic equivalence, briefly outlining possible future work on the topic.

2. Proof

Let us compute the expected value of the mutual information between the conditioning variable X and a generic variable Y_i:

$$E_i\left[\mathrm{I(X;Y_i)}\right] = \frac{1}{n}\sum_{i=1}^{n}\left(\mathrm{H(Y_i)} - \mathrm{H(Y_i|X)}\right) \tag{1}$$

where $\mathrm{I}(\cdot;\cdot)$ denotes the mutual information, $\mathrm{H}(\cdot)$ is the Shannon entropy and $\mathrm{H}(\cdot|\cdot)$ is the conditional entropy. Denoting by $\mathrm{C(Y_1, \ldots, Y_n)}$ the total correlation of the variables $Y_1, \ldots, Y_n$:

$$\begin{aligned}\mathrm{C(Y_1, \ldots, Y_n)} &= \sum_{y_1 \in \mathcal{Y}_1} \cdots \sum_{y_n \in \mathcal{Y}_n} p(y_1, \ldots, y_n) \log \frac{p(y_1, \ldots, y_n)}{p(y_1)p(y_2) \ldots p(y_n)} \\ &= \sum_{i=1}^{n} \mathrm{H(Y_i)} - \mathrm{H(Y_1, \ldots, Y_n)}, \end{aligned} \tag{2}$$

we can write

$$
\begin{aligned}
\mathrm{C}(\mathrm{Y_1}, \ldots, \mathrm{Y_n}) \quad + & \quad \mathrm{I}(\mathrm{X}; \mathrm{Y_1}, \ldots, \mathrm{Y_n}) = \\
= & \sum_{i=1}^{n} \mathrm{H}(\mathrm{Y_i}) - \mathrm{H}(\mathrm{Y_1}, \ldots, \mathrm{Y_n}) + \mathrm{H}(\mathrm{X}) + \mathrm{H}(\mathrm{Y_1}, \ldots, \mathrm{Y_n}) - \mathrm{H}(\mathrm{X}, \mathrm{Y_1}, \ldots, \mathrm{Y_n}) \\
= & \sum_{i=1}^{n} \mathrm{H}(\mathrm{Y_i}) - (\mathrm{H}(\mathrm{X}, \mathrm{Y_1}, \ldots, \mathrm{Y_n}) - \mathrm{H}(\mathrm{X})). \qquad (3)
\end{aligned}
$$

By using the conditional independence assumption $\mathrm{H}(\mathrm{Y_i}|\mathrm{Y_{i-1}}, \mathrm{Y_{i-2}}, \ldots, \mathrm{Y_1}, \mathrm{X}) = \mathrm{H}(\mathrm{Y_i}|\mathrm{X})$ the term $\mathrm{H}(\mathrm{X}, \mathrm{Y_1}, \ldots, \mathrm{Y_n})$ can be expanded recursively as:

$$
\begin{aligned}
\mathrm{H}(\mathrm{X}, \mathrm{Y_1}, \ldots, \mathrm{Y_n}) &= \mathrm{H}(\mathrm{Y_n}|\mathrm{Y_{n-1}}, \mathrm{Y_{n-2}}, \ldots, \mathrm{Y_1}, \mathrm{X}) + \mathrm{H}(\mathrm{Y_{n-1}}, \mathrm{Y_{n-2}}, \ldots, \mathrm{Y_1}, \mathrm{X}) \\
&= \mathrm{H}(\mathrm{Y_n}|\mathrm{X}) + \mathrm{H}(\mathrm{Y_{n-1}}, \mathrm{Y_{n-2}}, \ldots, \mathrm{Y_1}, \mathrm{X}) \\
&= \mathrm{H}(\mathrm{Y_n}|\mathrm{X}) + \mathrm{H}(\mathrm{Y_{n-1}}|\mathrm{X}) + \mathrm{H}(\mathrm{Y_{n-2}}, \mathrm{Y_{n-3}} \ldots, \mathrm{Y_1}, \mathrm{X}) \\
&= \ldots \\
&= \sum_{i=1}^{n} \mathrm{H}(\mathrm{Y_i}|\mathrm{X}) + \mathrm{H}(\mathrm{X}). \qquad (4)
\end{aligned}
$$

Substituting 4 into 3 yields

$$
\mathrm{C}(\mathrm{Y_1}, \ldots, \mathrm{Y_n}) + \mathrm{I}(\mathrm{X}; \mathrm{Y_1}, \ldots, \mathrm{Y_n}) = \sum_{\mathrm{i}=1}^{\mathrm{n}} (\mathrm{H}(\mathrm{Y_i}) - \mathrm{H}(\mathrm{Y_i}|\mathrm{X})) \qquad (5)
$$

from which 1 can be rewritten as

$$
E_i\left[\mathrm{I}(\mathrm{X}; \mathrm{Y_i})\right] = \frac{\mathrm{C}(\mathrm{Y_1}, \ldots, \mathrm{Y_n})}{n} + \frac{\mathrm{I}(\mathrm{X}; \mathrm{Y_1}, \ldots, \mathrm{Y_n})}{n}. \qquad (6)
$$

From the relation

$$
0 \leq \mathrm{I}(\mathrm{X}; \mathrm{Y_1}, \ldots, \mathrm{Y_n}) \leq \mathrm{H}(\mathrm{X}) \qquad (7)
$$

and the fact that $\mathrm{H}(\mathrm{X})$ is independent from n we have

$$
\lim_{n \to \infty} \frac{\mathrm{I}(\mathrm{X}; \mathrm{Y_1}, \ldots, \mathrm{Y_n})}{n} = 0 \qquad (8)
$$

and therefore

$$
\lim_{n \to \infty} E_i\left[\mathrm{I}(\mathrm{X}; \mathrm{Y_i})\right] = \frac{\mathrm{C}(\mathrm{Y_1}, \ldots, \mathrm{Y_n})}{n}. \qquad (9)
$$

3. Examples

As briefly explained in the introduction, studying approximations of the input-output mutual information in presence of stochastic resonance has great practical importance. Furthermore, in presence of stochastic resonance, the mutual information varies under one simple parameter, noise, and its variation assumes a well known bell shaped curve. It appears thus a very appealing setting for studying the relationship between spurious correlation and mutual information proved in the previous section. The following subsections present numerical simulations of two neuron models commonly used in the study of stochastic resonance, namely the binary threshold model and the FitzHugh-Nagumo neuron.

3.1. Binary threshold model

The most basic model for studying the stochastic resonance effect is given by the binary neuron threshold model [12], defined by the following equation:

$$R(t) = \begin{cases} 1 & if \quad S(t) + \xi(t) > \theta \\ 0 & otherwise \end{cases} \tag{10}$$

where $S(t)$ is the input, $R(t)$ is the output, $\xi(t)$ is Gaussian white random noise and θ is the threshold, which was set to 1 in our experiments. Two inputs were considered. The first is a periodic one, in detail $\frac{1}{2}(1 + \sin(2\pi f t))$ with $f = 10$Hz. The second is a classic aperiodic test signal [34, 35], generated by the convolution of Gaussian white noise with a Hanning window of width of 6 seconds, scaled to be underthreshold.

The results for the two input signals are presented in Fig.1 and Fig.2, respectively. Each figure reports, for various noise intensities, the average value of the mutual information between the input and the output of the units (which are, in our experiment, identical but conditionally independent). The plots also report the approximation obtained by dividing by n the total correlation computed over n units, for different values of n. From the plots it can be clearly seen that the estimation approaches the mutual information from underneath by increasing the number of units. Furthermore, the results suggest that even for a relatively limited number of neurons, the proposed approximation can be used for inferring the level of the average stochastic dependence between the conditioning input and the output.

3.2. FitzHugh-Nagumo model

The second model of neuron on which the relationship between mutual information and total correlation was tested is the FitzHugh-Nagumo model. This model constitutes a simplification of the Hodgkin-Huxley equations of electrical activity in the squid axon, and provides a good trade-off between simplicity and capacity of modeling the firing dynamics of real neurons.

Its dynamics can be formulated as follows:

$$\begin{aligned} \epsilon\dot{v} &= v(v-a)(1-v) - w + A + S(t) + \xi(t) \quad (11)\\ \dot{w} &= v - w - b \quad (12) \end{aligned}$$

where $v(t)$ and $w(t)$ are the state variables corresponding to fast and slow dynamics of the system, respectively, and $S(t)$ denotes again the input signal while $\xi(t)$ represents additive noise. The remaining terms are constant parameters which were set to commonly used values [34], namely $\epsilon = 0.005$, $a = 0.5$, $b = 0.15$, $A = 0.04$. The dynamics was simulated using the fourth-order Runge-Kutta method with a time step $\Delta t = 0.005$(sec), for 2^{14} simulation steps.

The input signal $S(t)$ is the same aperiodic signal introduced in the previous section. The output signal $R(t)$ was defined as the average number of spikes per second produced by the model. A spike is assumed to occur every time $v(t)$ crosses the value of 0.5 with a positive slope, and the averaging is computed by passing a 6 seconds unit-area symmetric Hanning window on the spike train.

Fig.3 shows the result depicting the expected value of the mutual information between the input signal $S(t)$ and the output signal $R(t)$, and the total correlation among two and three neurons. From the figure, we note that the approximation of the mutual information in terms of total correlation remains valid even for neurons with internal dynamics as the FitzHugh-Nagumo model.

4. Conclusions and future works

In this paper, we show that if multiple, observable, random variables are directly influenced by the same latent random variable, then the value of the spurious correlation among the observable variables can be used for inferring the strength of relationship between the latent variable and the observable variables. In particular, this work proves that the spurious correlation, expressed as the total correlation among the observable variables divided by the number of variables, is asymptotically equal to the average mutual information between the latent variables and each of the observable variables, as the number of variables tends to infinity. The result finds practical application in all the cases in which an unknown signal can be measured by multiple sources, and there are controllable parameters that can modulate the information transfer between the unobservable signal and the measured information.

As an example, this paper shows the possible application of the method to the tuning of the noise level in Stochastic Resonance. Stochastic resonance has been gathering considerable attention in engineering. Although the emergence of stochastic resonance was shown to occur for a wide range of noise intensities under opportune conditions [36], the maximization of the SR effect remains of fundamental interest in practical applications.

In literature, the emergence of SR is measured and evaluated by using both the input and the output signal of the system. However, from the perspective of practical applications, both in physics and engineering, the necessity of knowing the input for maximizing the stochastic resonance effect, i.e. for actually being able to better observe such input, remains a great obstacle in the full exploitation of SR. For instance, in the case of a subthreshold signal, the threshold could be actually the minimum level measurable with the available instrumentation. In such cases, SR may be the only way for obtaining information on the signal, and thus assuming the knowledge of the input for maximizing the SR effect makes no sense.

This paper aims at filling this lack and opening the possibility of a more widespread application of SR in engineering applications. We must note that this paper only proves that for a sufficient large number of observable variables, their total correlation divided by the number of variables asymptotically tends to the mutual information. This result does not guarantee that the level of noise that maximizes the total correlation is the level of noise that maximizes the mutual information for a finite number of variables. However, the numerical results presented in the paper show that, in practice, even for a relatively low number of units the total correlation can be used to identify the optimal noise with a good approximation. Future work will need to concentrate in identifying the conditions under which this relationship is satisfied.

The results presented in this paper turn out to be interesting for the biological field as well. In particular, they suggests that neural populations may be able to tune the SR effect using only total correlation between neurons of the same layer, without the need to know the actual value of the original stimulus. Even more interestingly, the link between total correlation and mutual information indicates the possibility that total correlation may be used for training complex networks, in the same that way mutual information is used in Infomax [37] for the spontaneous development of perceptual network layers. In this regard, the recent development of techniques for computing approximations of the joint probability of a high number of continuous variables [38], assumes particular importance, and finding a biologically plausible way in which these approaches may take place in neural networks surely deserves further investigation.

5. Acknowledgments

This work was supported by a Grant-in-Aid for Young Scientists (B) Grant Number 24700193.

[1] L. Gammaitoni, P. Hänggi, P. Jung, F. Marchesoni, Stochastic resonance, Reviews of Modern Physics 70 (1) (1998) 223–287.

[2] R. Benzi, A. Sutera, A. Vulpiani, The mechanism of stochastic resonance, J. of Physics A: mathematical and general 14 (1981) 453–457.

[3] F. Moss, L. M. Ward, W. G. Sannita, Stochastic resonance and sensory information processing: a tutorial and review of application, Clinical Neurophysiology 115 (2) (2004) 267 – 281.

[4] A. A. Priplata, B. L. Patritti, J. B. Niemi, R. Hughes, D. C. Gravelle, L. A. Lipsitz, A. Veves, J. Stein, P. Bonato, J. J. Collins, Noise-enhanced balance control in patients with diabetes and patients with stroke, Annals of Neurology 59 (1) (2006) 4–12.

[5] G. Finocchio, I. Krivorotov, X. Cheng, L. Torres, B. Azzerboni, Micromagnetic understanding of stochastic resonance driven by spin-transfertorque, Physical Review B 83 (13) (2011) 134402.

[6] B. Franzke, B. Kosko, Noise can speed convergence in markov chains, Phys. Rev. E 84 (2011) 041112.

[7] V. Hari, G. Anand, A. Premkumar, A. Madhukumar, Design and performance analysis of a signal detector based on suprathreshold stochastic resonance, Signal Processing 92 (7) (2012) 1745 – 1757.

[8] T. Oya, Thermal-noise-exploiting operations of single-electron majority logic circuits with conventional clock signals, Nanotechnology, IEEE Transactions on 11 (1) (2012) 134–138.

[9] T. Hindo, S. Chakrabartty, Noise-exploitation and adaptation in neuromorphic sensors, in: Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8339 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 2012.

[10] Q. Li, Z. Li, J. Shen, R. Gao, A novel spectrum sensing method in cognitive radio based on suprathreshold stochastic resonance, in: Communications (ICC), 2012 IEEE Int.Conf. on, 2012, pp. 4426–4430.

[11] A. Das, N. G. Stocks, E. L. Hines, Enhanced coding for exponentially distributed signals using suprathreshold stochastic resonance, Communications in Nonlinear Science and Numerical Simulation 14 (1) (2009) 223 – 232.

[12] N. G. Stocks, Suprathreshold stochastic resonance in multilevel threshold systems, Phys. Rev. Lett. 84 (2000) 2310–2313.

[13] G. Deco, B. Schrmann, Stochastic resonance in the mutual information between input and output spike trains of noisy central neurons, Physica D: Nonlinear Phenomena 117 (14) (1998) 276 – 282.

[14] T. Kondo, T. Munakata, Stochastic resonance and self-tuning: A simple threshold system, Phys. Rev. E 79 (2009) 061121.

[15] N. Brunel, J.-P. Nadal, Mutual information, fisher information, and population coding, Neural Comput. 10 (7) (1998) 1731–1757.

[16] T. Hoch, G. Wenning, K. Obermayer, Optimal noise-aided signal transmission through populations of neurons, Physical Review E 68 (1) (2003) 011911.

[17] M. D. McDonnell, N. G. Stocks, Maximally informative stimuli and tuning curves for sigmoidal rate-coding neurons and populations, Phys. Rev. Lett. 101 (2008) 058103.

[18] G. Wenning, K. Obermayer, Activity driven adaptive stochastic resonance, Phys. Rev. Lett. 90 (2003) 120602.

[19] Y. Guo, J. Tan, Suprathreshold stochastic resonance in multilevel threshold system driven by multiplicative and additive noises, Communications in Nonlinear Science and Numerical Simulation 18 (10) (2013) 2852 – 2858.

[20] S. Blanchard, D. Rousseau, F. Chapeau-Blondeau, Noise enhancement of signal transduction by parallel arrays of nonlinear neurons with threshold and saturation, Neurocomputing 71 (13) (2007) 333 – 341.

[21] A. P. Georgopoulos, A. B. Schwartz, R. E. Kettner, Neuronal population coding of movement direction, Science 233 (4771) (1986) 1416–1419.

[22] A. Pouget, P. Dayan, R. Zemel, Information processing with population codes, Nature Reviews Neuroscience 1 (2) (2000) 125–132.

[23] S. B. Laughlin, R. R. de Ruyter van Steveninck, J. C. Anderson, The metabolic cost of neural information, Nature Neuroscience 1 (1) (1998) 36–41.

[24] K. Pearson, Mathematical contributions to the theory of evolution.–on a form of spurious correlation which may arise when indices are used in the measurement of organs, Proceedings of the Royal Society of London 60 (1896) 489–498.

[25] J. Aldrich, Correlations genuine and spurious in pearson and yule, Statistical Science 10 (4) (1995) 364–376.

[26] G. U. Yule, On the interpretation of correlations between indices or ratios, Journal of the Royal Statistical Society 73 (6/7) (1910) 644–647.

[27] G. U. Yule, An Introduction to the Theory of Statistics, 6th ed., Griffin, London, 1922.

[28] H. A. Simon, Spurious Correlation: A Causal Interpretation, Journal of the American Statistical Association 49 (267) (1954) 467–479.

[29] W. McGill, Multivariate information transmission, Information Theory, Transactions of the IRE Professional Group on 4 (4) (1954) 93–111.

[30] S. Watanabe, Information theoretical analysis of multivariate correlation, IBM Journal of Research and Development 4 (1) (Jan.) 66–82.

[31] W. R. Garner, Uncertainty and Structure as Psychological Concepts, JohnWiley & Sons, New York, 1962.

[32] T. S. Han, Nonnegative entropy measures of multivariate symmetric correlations, Information and Control 36 (2) (1978) 133 – 156.

[33] M. Studeny, Y, J. Vejnarová, The multiinformation function as a tool for measuring stochastic dependence, in: Learning in Graphical Models, 1998, pp. 261–298.

[34] J. J. Collins, C. C. Chow, T. T. Imhoff, Aperiodic stochastic resonance in excitable systems, Phys. Rev. E 52 (1995) R3321–R3324.

[35] D. Nozaki, Y. Yamamoto, Enhancement of stochastic resonance in a fitzhugh-nagumo neuronal model driven by colored noise, Physics Letters A 243 (56) (1998) 281 – 287.

[36] J. J. Collins, C. C. Chow, T. T. Imhoff, Stochastic resonance without tuning, Nature 376 (6537) (1995) 236–238.

[37] R. Linsker, Self-organization in a perceptual network, Computer 21 (3) (1988) 105–117.

[38] L. Faivishevsky, J. Goldberger, Ica based on a smooth estimation of the differential entropy, in: Neural Information Processing Systems (NIPS), 2008, pp. 433–440.

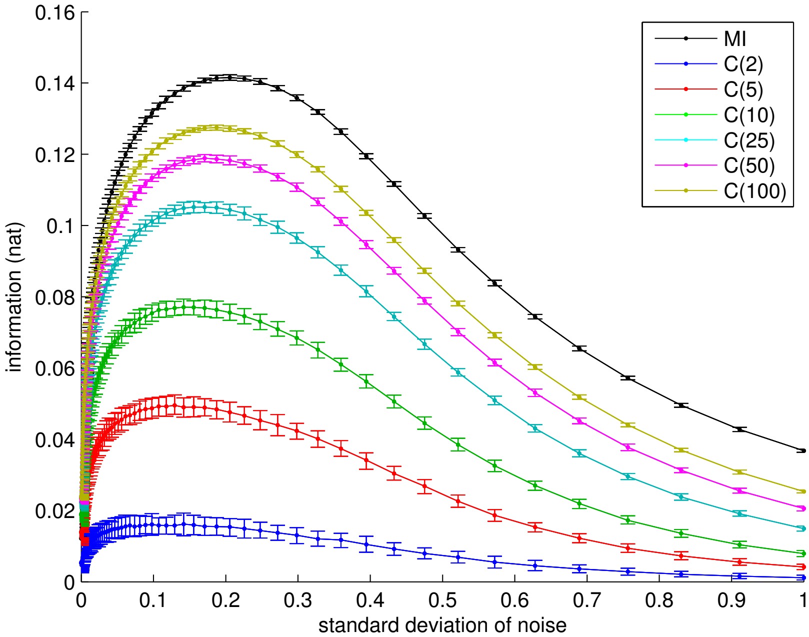

Figure 1: The expected value of the mutual information and the total correlation approximation for different noise levels obtained using the periodic signal reported in the text. The top curve corresponds to the mutual information while the other curves correspond, from bottom to top, to the total correlation approximation obtained for 2, 5, 10, 25, 50, and 100 units, respectively. In this and in the following figures, for each setting, the simulation was repeated 100 times, and the 95% confidence interval, here indicated by the error bars, was computed. The asymptotic convergence of the total correlation (computed only from the output values) to the mutual information can be clearly observed in the plot.

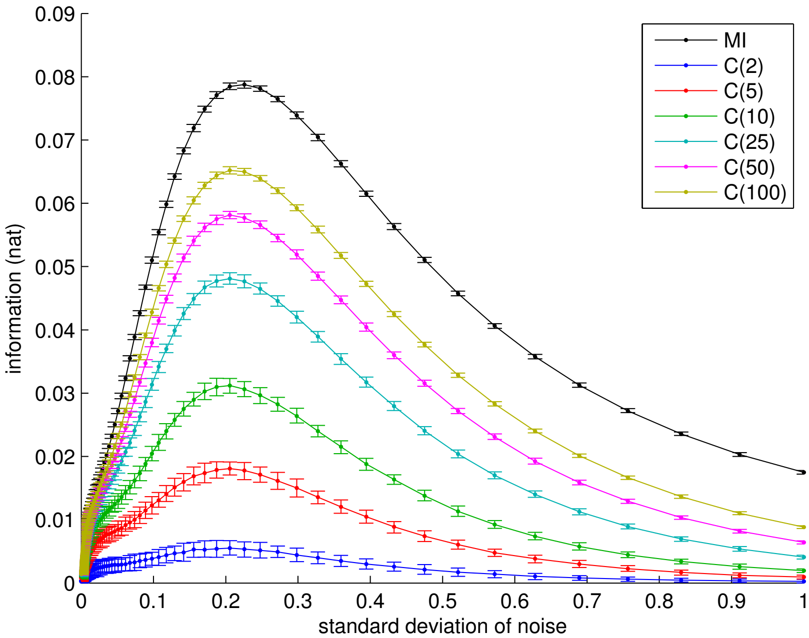

Figure 2: The expected value of the mutual information and the total correlation approximation for different levels of noise plotted for different number of units obtained for the aperiodic input signal described in the text.

.

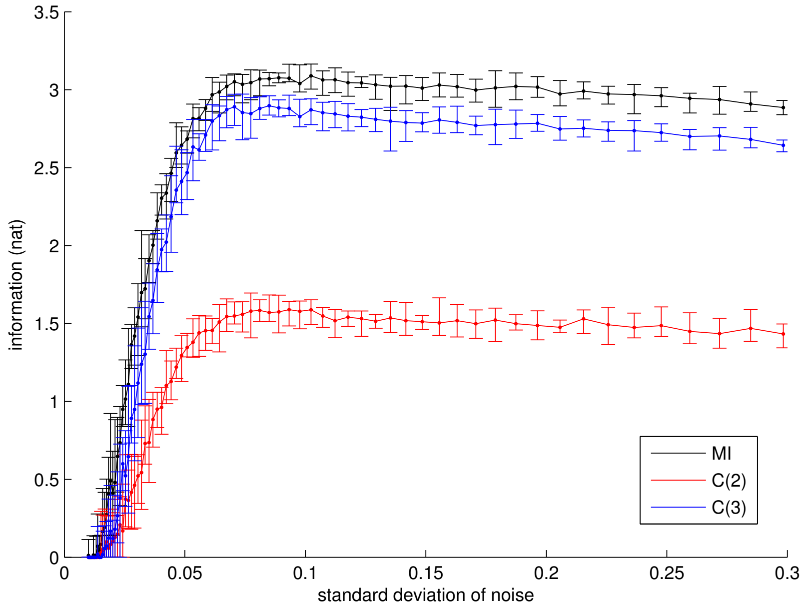

Figure 3: The expected values of the mutual information (top curve) and the total correlation among two (bottom curve) and three (central curve) FitzHugh-Nagumo neurons computed for different noise settings. It is easy to see that the total correlation among the units, computed only from output signals, gives an approximation of the mutual information between the input and the outputs. As a simple discretization of the output in 500 bins was taken to compute the (continuous) total correlation between the unit outputs $R(t)$, due to memory limitations only experiments for 2 and 3 units are reported.