

# Nanosignal Processing: Stochastic Resonance in Carbon Nanotubes That Detect Subthreshold Signals

Ian Y. Lee, Xiaolei Liu, Bart Kosko,\* and Chongwu Zhou

Electrical Engineering Department, University of Southern California,  
Los Angeles, California 90089-2564

Received September 25, 2003; Revised Manuscript Received October 11, 2003

## ABSTRACT

Experiments confirm that small amounts of noise help a nanotube transistor detect noisy subthreshold electrical signals. Gaussian, uniform, and impulsive (Cauchy) noise produced this feedforward stochastic-resonance effect by increasing both the nanotube system's mutual information and its input–output correlation. The noise corrupted a synchronous Bernoulli or random digital sequence that fed into the thresholdlike nanotube transistor and produced a Bernoulli sequence. Both Shannon's mutual information and correlation measured the performance gain by comparing the input and output sequences. This nanotube SR effect was robust: it persisted even when infinite-variance Cauchy noise corrupted the signal stream. Such noise-enhanced signal processing at the nanolevel promises applications to signal detection in wideband communication systems and biological and artificial neural networks.

Noise can help carbon nanotube transistors detect subthreshold electrical signals by increasing the transistor's input–output mutual information or correlation. Several researchers have demonstrated the stochastic resonance (SR) effect for various types of threshold units or neurons.<sup>1–6</sup> Experiments with p-type nanotube transistors confirmed the specific SR prediction based on the theoretical finding that simple memoryless threshold neurons exhibit SR for almost all finite-variance and infinite-variance noise types.<sup>7</sup> The experiments used three types of additive noise (Gaussian, uniform, and infinite-variance<sup>1</sup> Cauchy noise) and different combinations of subthreshold ON/OFF electrical signals. Figure 1 shows the nonmonotonic signature of SR for white Gaussian noise and the thresholdlike nonlinearity of the nanotube transistors.<sup>8–13</sup> The modes of the mutual-information and correlation curves occurred for nonzero noise strength with a standard deviation of at least 0.01.

The nanotube experiments produced the SR effect for both the Shannon mutual information and the input–output correlation<sup>14</sup> of noisy Bernoulli sequences. The mutual information  $I(S, Y)$  subtracts the noisy channel's (the transistor's) output conditional entropy  $H(Y|S)$  from its unconditional entropy  $H(Y)$ :  $I(S, Y) = H(Y) - H(Y|S)$ . The input signal  $S$  was a random binary voltage that produced a random output  $Y$  in the form of a transistor current. The

correlation measure found the normalized zero-lag cross-correlation

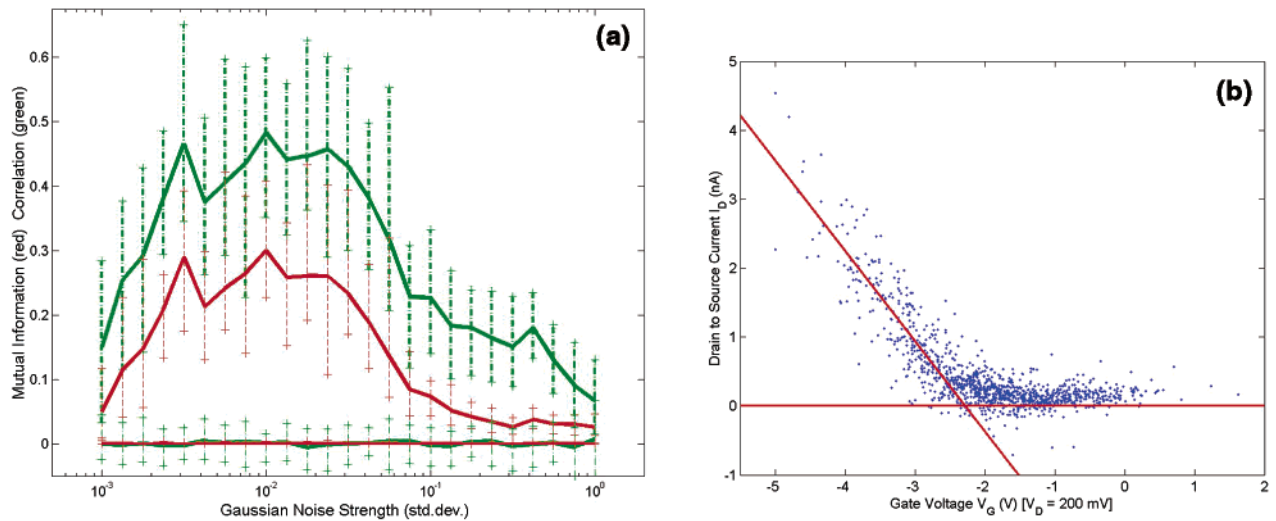
$$r_{SY}(l) = \frac{1}{N} \sum_{k=1}^N s(k) y(k-l)$$

of the two sequences with subtracted means. The measures did not assume that the nanotube detector had a special structure and did not impose a threshold scheme on the experiment.

Figure 1b shows the thresholdlike nonlinearity of the nanotube transistor in response to the noisy input signal. The transconductance  $G$  related the output drain-to-source current  $I$  to the input gate voltage  $V$  and the threshold voltage  $V_T$  in a memoryless signal function:  $I = G(V - V_T)$  if  $V \leq V_T$  and zero otherwise. We note that the threshold neuron model lacks the internal state dynamics of the FitzHugh–Nagumo (FHN) model.<sup>15</sup> The transconductance  $G$  was negative because the pristine (undoped) nanotube transistors exhibited current–voltage characteristics that were consistent with p-type transistors. Linear regression extrapolated the nonlinearity and estimated the threshold voltage.

Each of the nanotube experiments (Supporting Information) applied 32 independent trials of 1000-symbol input sequences for 24 noise levels per type and over a range of gate voltages. The 24 sampled noise levels ranged from 0.001 to 1 standard deviation (dispersion for infinite-variance Cauchy) linearly on a logarithmic scale. The noisy input was

\* Corresponding author. E-mail: kosko@sipi.usc.edu.



**Figure 1.** (a) Stochastic resonance with additive white Gaussian noise. The CNT-FET detector’s mutual information (top red curve) and zero-lag correlation (top green curve) increase for small amounts of noise and then decrease for larger amounts. The control experiments gave the flat non-SR mutual information (bottom red curve) and correlation (bottom green curve) when no nanotube bridged the source and drain electrodes. The SR mode or optimal noise level had the same standard deviation value of 0.01 for both performance measures. Each vertical dashed bar occurs at 1 of the 24 sampled noise values and shows the maximum and the minimum range of 32 averaged experimental trials. The solid polygonal line connects the means of those 24 sets of experiments. The random input sequence  $S$  was a Bernoulli sequence of ON/OFF values with additive white Gaussian noise. The random sequence  $Y$  was the output of the nanotube threshold detector. Shown is one of four such successful combinations of input binary values with the parameter choices ON =  $-1.6$  V and OFF =  $-1.4$  V. Each trial applied 1000 subthreshold symbols to the detector. The input signal was the analog voltage representation of the symbol sequence  $S$  at approximately 10-ms intervals. The output signal was the nanotube current. The data acquisition measured and averaged 10 samples at 100 ksamples/s near the end of each symbol interval to estimate the output sequence (Supporting Information). A  $\chi^2$  test and a Kolmogorov–Smirnov test both rejected the similarity between a monotonically decreasing  $\beta$  probability density function and the two SR curves ( $p < 0.001$ ). (b) Thresholdlike (nonlinear) gate effect of the p-type CNT-FET detector. Each point shows the detector’s response to one random input symbol. The experimental data showed that the CNT-FET detector behaved as a threshold in response to the noisy input signal stream. The gate effect showed little hysteresis. This differed from the hysteretic curve that a semiconductor parameter analyzer captured from the detector (Supporting Information) and differed from the typical hysteretic loops in ref 18. Linear regression gave an approximate threshold gate voltage of  $V_T = -2.3$  V ( $\beta_0 = -2.99$  nA,  $\beta_1 = -1.31$  nA/V,  $p$  value  $< 0.0001$ ) for the transistor current equation  $I = G(V - V_T)$  if  $V \leq V_T$  and zero otherwise.

a synchronized Bernoulli sequence of independent random (subthreshold) ON/OFF values and additive white noise of three types. So there was no timing noise in the pulse train as in the FHN neuron model.<sup>16</sup> The discrete-time noise was white because the noise samples were uncorrelated in time. So the discrete-time Fourier transform was  $2\pi$ -periodic and produced a flat noise power spectrum over the interval  $[0, 2\pi]$ .<sup>17,18</sup> Synchronization allows the nanotube systems to implement a variety of algorithms from signal processing and communications.

The ON/OFF values in Figure 1a were ON =  $-1.6$  V and OFF =  $-1.4$  V. The input updated the symbols about once every 10 ms. A 200-mV drain-source voltage biased the nanotube at room temperature in vacuum. The experiment measured and averaged 10 samples of the detector output at 100 ksymbols/s near the end of each symbol interval to estimate the output sequence.

A histogram of the output sequence gave the discrete probability density function  $P(Y = Y_i) = p_i$  that computed the unconditional Shannon entropy:

$$H(Y) = - \sum_{i=1}^N p_i \ln p_i \quad (1)$$

for mutual information without converting the detector output

into a binary sequence with a threshold scheme. Sorting the output sequence based on the input symbol and then applying the histogram gave the conditional output discrete probability density function  $P_{Y_i|S}(Y = Y_i|S = S_j) = p_{ji}/p_j$  conditioned on the input symbols that computed the conditional entropy:

$$H(Y|S) = - \sum_{i=1}^N \sum_{j=1}^N p_{ji} \ln \left( \frac{p_{ji}}{p_j} \right) \quad (2)$$

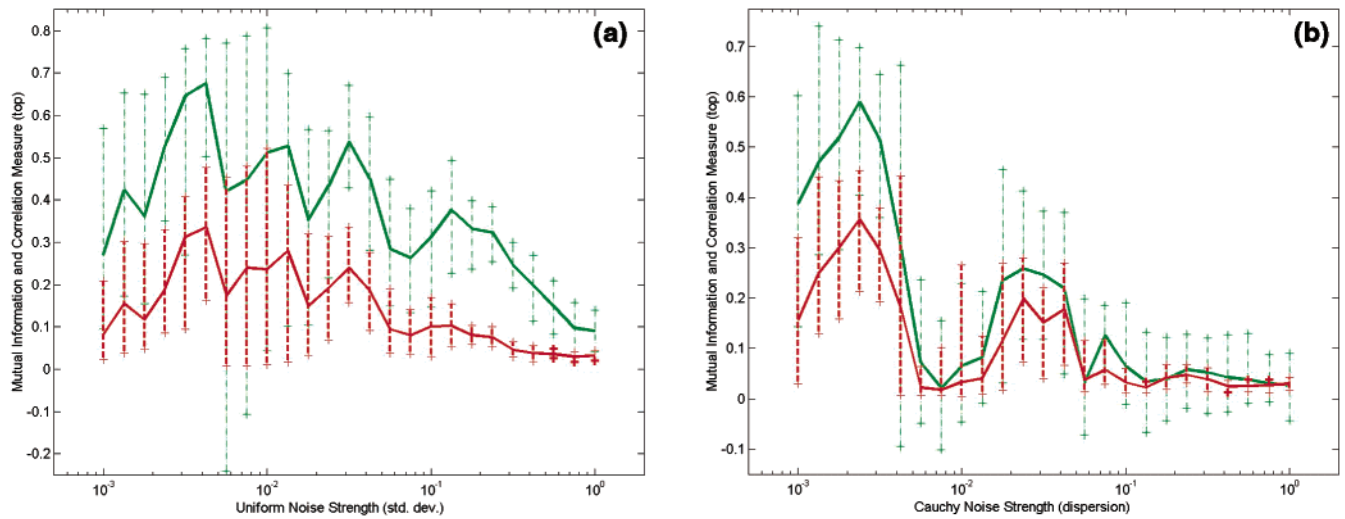
The mutual information measure was the difference between the unconditional and conditional entropies:

$$I(S, Y) = H(Y) - H(Y|S) \quad (3)$$

Cross correlation compared the input and the output symbol sequences and gave a scalar representation with its zero-lag value:

$$r_{SY}(0) = \sum_{k=1}^N s(k) y(k) \quad (4)$$

Converting the input Bernoulli sequence to bipolar form (mapping ON to  $+1$  and OFF to  $-1$ ) made it approximately



**Figure 2.** (a) Stochastic resonance with additive white uniform noise. All four combinations of input voltage values produced a clear SR response in both mutual information (bottom red curve) and input–output correlation (top green curve) just as with additive white Gaussian noise. Shown is the SR effect for the subthreshold signal ON = −1.8 V and OFF = −1.6 V. The SR mode is at 0.04 standard deviation. (b). Robust stochastic resonance with additive white Cauchy noise. This highly impulsive noise has infinite variance and infinite higher-order moments. The Cauchy-noise experiment produced a measurable SR effect for two of the four combinations of input voltages. Shown is an approximate SR effect for the subthreshold signal ON = −2 V and OFF = −1.8 V. The SR mode lies at about the 0.003 dispersion value. Several SR researchers have found multiple modes in the plot of system performance against noise strength.<sup>51–53</sup> The limited dynamic range [−5V, 5V] of the data acquisition equipment (Supporting Information) may have produced the second peak in the graph as a truncation artifact because it clipped large spikes when it realized the infinite-variance Cauchy noise. The clipping affected more than 0.1% of the noise only for dispersions greater than 0.01.

zero mean (equal numbers of +1’s and −1’s give exactly zero mean) and noise-free. Subtracting the sample mean from the output sequence improved the match between similar input and output sequences. A normalization scheme gave the normalized correlation measure:<sup>14</sup>

$$C(S, Y) = \frac{\sum_{k=1}^N s(k) y(k)}{\sqrt{\sum_{k=1}^N s(k) s(k)} \sqrt{\sum_{k=1}^N y(k) y(k)}} \quad (5)$$

It divided the zero-lag cross correlation  $r_{sY}(0)$  by the square root of the energy of the input and the output sequences where the energy of a sequence is the same as the zero-lag value of its autocorrelation:

$$|x| = \sum_{k=1}^N x^2(k) = \sum_{k=1}^N x(k) x(k - l)|_{l=0} = r_{XX}(0) \quad (6)$$

Nanotube field-effect transistor technology produced detectors that could exhibit hysteresis<sup>19–21</sup> or react to adsorbed molecules.<sup>22–24</sup> The experiment applied subthreshold symbols that were at least two standard deviations away from the far leg of the hysteretic loop. The effective  $I-V_G$  curve in Figure 1b, as collected from the detector response to the input symbols, showed that the experiment produced evidence of the SR effect despite the potential hysteretic effect.

The experiment found the SR effect for mutual information and correlation for Gaussian and uniform noise and for four

combinations of binary symbols: (−2.0, −1.8), (−1.8, −1.6), (−1.6, −1.4), and (−1.4, −1.2) V. Figure 1a shows the SR effect for additive white Gaussian noise and the subthreshold signal pair ON = −1.6 V and OFF = −1.4 V. The SR mode of the mutual-information curve is 6 times the value at minimal noise. The SR mode of the correlation curve is 3 times the value at minimal noise. Figure 2a shows the SR effect for additive white uniform noise and the signal pair ON = −1.8 V and OFF = −1.6 V.

We also passed impulsive or infinite-variance white noise through the nanotube detector to test whether it was robust to occasional large noise spikes. We chose the highly impulsive Cauchy noise<sup>1</sup> for this task. This infinite-variance noise probability density function had the form

$$p(n) = \frac{\gamma}{\pi(n^2 + \gamma^2)}$$

for zero location and finite dispersion  $\gamma$ . Figure 2b shows that a diminished SR effect still persists for Cauchy noise with the subthreshold signal pair ON = −2.0 V and OFF = −1.8 V. Not all Cauchy experiments produced a measurable SR effect.

These SR results suggest that nanotubes can exploit noise in other signal-processing tasks if advances in nanotube device technology can overcome the problems of hysteresis and parasitic capacitance that affect logic circuits<sup>25</sup> and high-frequency signals.<sup>26</sup> The nanotube signal detectors might apply to broadband<sup>27,28</sup> or optical communication systems<sup>29</sup> that use submicroamp currents and attenuated signals in noise because our nanotube detectors used nanoamp current and could distinguish between subthreshold binary symbols. The

detectors might apply to parallel signal processing<sup>30</sup> at the nanolevel because they could have a small minimum feature size<sup>31</sup> in vast parallel arrays of nanotubes. The parallel detectors could apply to spread spectrum communications: each nanotube can act as an antenna<sup>32</sup> that matches a separate frequency channel<sup>33</sup> in frequency hopping and perhaps in other types of spread spectrum communications.<sup>34</sup> A nanotube's length can code for a given frequency<sup>35</sup> while chemical adsorption can tune a nanotube's threshold.<sup>23,24</sup> The detectors might apply to chemical detection and parallel field programming by tuning the threshold chemically. The nanotube detectors can also operate in a biological environment such as saline solution.<sup>36</sup> The nanotube detectors could interface with biological systems because an electrolyte can act as their gate.<sup>36,37</sup> The nanotube detectors might also help implement pulse-train neural networks and exploit noise in biological<sup>38-49</sup> or robotic systems because the detectors are threshold devices similar to spiking neurons.<sup>50</sup>

**Acknowledgment.** The National Science Foundation (grant no. ECS-0070284) supported this research. We thank Chao Li, Sanya Mitaim, and Ashok Patel for helpful discussions.

**Supporting Information Available:** Methods and Materials. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## References

- (1) Kosko, B.; Mitaim, S. *Neural Networks* **2003**, *16*, 755.
- (2) Kosko, B.; Mitaim, S. *IJCNN'02 IEEE Proc. Int. Joint Conf. Neural Networks*; **2002**, *2*, 1980.
- (3) Moss, F.; Pierson, D.; O'Gorman, D. *Int. J. Bifurc. Chaos* **1994**, *6*, 1383-1397.
- (4) Gingl, Z.; Kiss, L.; Moss, F. *Europhys. Lett.* **1995**, *29*, 191.
- (5) Inchirosa, M. E.; Robinson, J. W. C.; Bulsara, A. R. *Phys. Rev. Lett.* **2000**, *85*, 3369-3372.
- (6) Gammaitoni, L.; Bulsara, A. R. *Phys. Rev. Lett.* **2002**, *88*, 230601.
- (7) Kosko, B.; Mitaim, S. *Phys. Rev. E* **2001**, *64*, 051110.
- (8) Tans, S. J.; Verschueren, R. M.; Dekker, C. *Nature* **1998**, *393*, 49.
- (9) Zhou, C.; Kong, J.; Dai, H. *Appl. Phys. Lett.* **2000**, *76*, 1597.
- (10) Martel, R.; Schmidt, T.; Shea, H. R.; Hertel, T.; Avouris Ph. *Appl. Phys. Lett.* **1998**, *73*, 2447.
- (11) Cheung, C. L.; Kurtz, A.; Park, H.; Lieber, C. M. *J. Phys. Chem. B* **2002**, *106*, 2429.
- (12) Liu, J.; Rinzler, A. G.; Dai, H.; Hafner, J. H.; Bradley, R. K.; Boul, P. J.; Lu, A.; Iverson, T.; Shelimov, K.; Huffman, C. B.; Rodriguez-Macias, F.; Shon, Y. S.; Lee, T. R.; Colbert, D. T.; Smalley, R. E. *Science* **1998**, *280*, 1253-1256.
- (13) Saito, R.; Fujita, M.; Dresselhaus, G.; Dresselhaus, M. S. *Appl. Phys. Lett.* **1992**, *60*, 2204.
- (14) Collins, J. J.; Chow, C. C.; Capela, A. C.; Imhoff, T. T. *Phys. Rev. E* **1996**, *54*, 5575.
- (15) Pei, X.; Bachmann, K.; Moss, F. *Phys. Lett. A*, **1995**, *206*, 61-65.
- (16) Pei, X.; Wilkens, L.; Moss, F. *Phys. Rev. Lett.* **1996**, *77*, 4679-4682.
- (17) Oppenheim, A. V.; Schaffer, R. W.; Buck, J. R. *Discrete-Time Signal Processing*, 2nd ed.; Prentice-Hall: Upper Saddle River, NJ, 1999.
- (18) Ingle, V. K.; Manolakis, D. G.; Kogon, S. *Statistical and Adaptive Signal Processing: Spectral Estimation, Signal Modeling, Adaptive Filtering and Array Processing*; McGraw-Hill: New York, 1999.
- (19) Kim, W.; Javey, A.; Vermesh, O.; Wang, Q.; Li, Y.; Dai, H. *Nano Lett.* **2003**, *3*, 193.
- (20) Fuhrer, M. S.; Kim, B. M.; Durkop, T.; Branglingler, T. *Nano Lett.* **2002**, *2*, 757.
- (21) Radosavljevic, M.; Freitag, M.; Thadani, K. V.; Johnson, A. T. *Nano Lett.* **2002**, *2*, 761.
- (22) Ong, K. G.; Zeng, K.; Grimes, C. A. *IEEE Sensors J.* **2002**, *2*, 82.
- (23) Collins, P. G.; Bradley, K.; Ishigami, M.; Zettl, A. *Science* **2000**, *287*, 1801.
- (24) Kong, J.; Franklin, N. R.; Zhou, C.; Chapline, M. G.; Peng, S.; Cho, K.; Dai, H. *Science* **2000**, *287*, 622-625.
- (25) Bachtold, A.; Hadley, P.; Nakanishi, T.; Dekker, C. *Science* **2001**, *294*, 1317.
- (26) Harris, D.; Horowitz, M. A. *IEEE J. Solid-State Circuits* **1997**, *32*, 1702-1711.
- (27) Bulsara, A. R.; Zador, A. *Phys. Rev. E* **1996**, *54*, R2185.
- (28) Barbay, S.; Giacomelli, G.; Marin, F. *Phys. Rev. E* **2000**, *63*, 051110-051118.
- (29) Chen, Y.-C.; Ravivakar, N. R.; Schadler, L. S.; Ajayan, P. M.; Zhao, Y.-P.; Lu, T.-M.; Wang, G.-C.; Zhang, X.-C. *Appl. Phys. Lett.* **2002**, *81*, 975-977.
- (30) Stocks, N. G. *Phys. Rev. E* **2001**, *63*, 041114-041122.
- (31) Rochefort, A.; Ventra, M. D.; Avouris, Ph. *Appl. Phys. Lett.* **2001**, *78*, 2521.
- (32) Liang, W.; Bockrath, M.; Bozovic, D.; Hafner, J. H.; Tinkham, M.; Park, H. *Nature* **2001**, *411*, 665-669.
- (33) Blake, L. V. *Antennas*, 2nd ed.; Artech House Inc.: Dedham, MA, 1984.
- (34) Imer R.; Fettweis, G. In *The 5th International Symposium on Wireless Personal Multimedia Communications*; IEEE: Washington, D.C., 2002; Vol. 2, Iss. 27-30, pp 412-416.
- (35) Huang, S.; Cai, X.; Liu, J. *J. Am. Chem. Soc.* **2003**, *125*, 5636-5637.
- (36) McEuen, P. L.; Fuhrer, M. S.; Park, H. *IEEE Trans. Nanotech.* **2002**, *1*, 78.
- (37) Kruger, M.; Buitelaar, M. R.; Nussbaumer, T.; Schonenberger, C.; Forro, L. *Appl. Phys. Lett.* **2001**, *78*, 1291.
- (38) Russel, D. F.; Wilkens, Lon. A.; Moss, F. *Nature* **1999**, *402*, 291.
- (39) Jaramillo, F.; Wiesenfeld, K. *Nat. Neurosci.* **1998**, *1*, 384.
- (40) Bulsara, A. R.; Boss, R. D.; Jacobs, E. W. *Biol. Cybern.* **1989**, *61*, 211-222.
- (41) Pantazelou, E.; Dames, C.; Moss, F.; Douglass, J.; Wilkens, L. *Int. J. Bifurc. Chaos* **1995**, *5*, 101-108.
- (42) Pei, X.; Wilkens, L.; Moss, F. *J. Neurophysiol.* **1996**, *76*, 3002-3011.
- (43) Douglass, J. K.; Wilkens, L.; Pantazelou, E.; Moss, F. *Nature* **1993**, *365*, 337-340.
- (44) Levin, J. E.; Miller, J. P. *Nature* **1996**, *380*, 165-168.
- (45) Miller, J. P.; Jacobs, G. A.; Theunissen, F. E. *J. Neurophysiol.* **1991**, *66*, 1680-1689.
- (46) Miller, J. P.; Jacobs, G. A.; Theunissen, F. E. *J. Neurophysiol.* **1991**, *66*, 1690-1703.
- (47) Braun, H. A.; Wissing, H.; Schäfer, K.; Hirsch, M. C. *Nature* **1994**, *367*, 270-273.
- (48) Gluckman, B. J.; Netoff, T. I.; Neel, E. J.; Ditto, W. L.; Spano, M. L.; Schiff, S. J. *Phys. Rev. Lett.* **1996**, *77*, 4098-4101.
- (49) Collins, J. J.; Imhoff, T. T.; Grigg, P. *J. Neurophysiol.* **1996**, *76*, 642-645.
- (50) Rieke, F.; Warland, D.; Steveninck, RdRv. *Spikes: Exploring the Neural Code*; MIT Press: Cambridge, MA, 1999.
- (51) Vilar, J. M. G.; Rubí, J. M. *Phys. Rev. Lett.* **1997**, *78*, 2882-2885.
- (52) Lindner, J. F.; Breen, B. J.; Wills, M. E.; Bulsara, A. R.; Ditto, W. L. *Phys. Rev. E* **2001**, *63*, 051107.
- (53) Matyjaskiewicz, S.; Krawiecki, A.; Holyst, J. A.; Kacperski, K.; Ebeling, W. *Phys. Rev. E* **2001**, *63*, 026215.

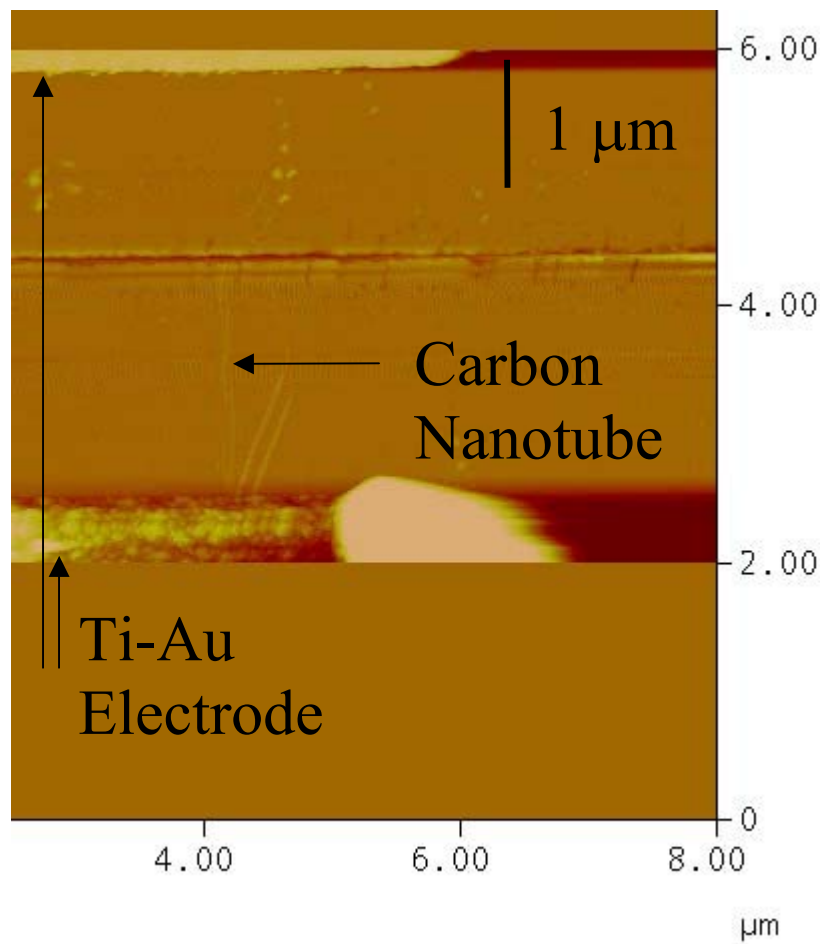
NL0348239

## SUPPORTING INFORMATION

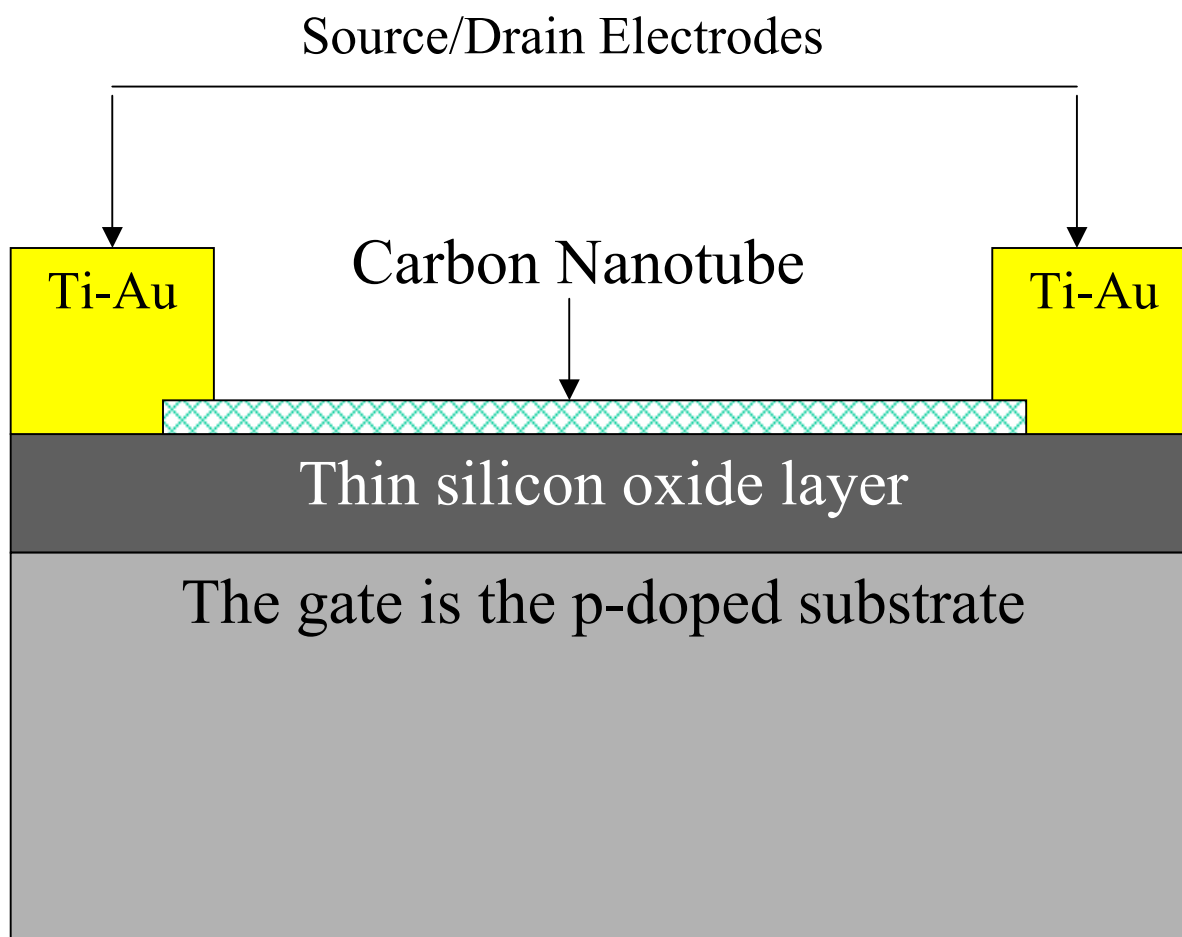
### **Materials and Methods**

The experiment tested a p-type carbon nanotube field effect transistor (CNT-FET) as a threshold detector with subthreshold input signal plus noise. The input signal was a Bernoulli random variable corrupted by a zero-mean Gaussian, uniform, or zero-location Cauchy distributed noise. The experiment captured the detector output signal in response to the noisy input and compared against the Bernoulli input signal to determine the detector performance. The experiment sought evidence of the stochastic resonance (SR) effect: detection that improved with increasing noise strength before deteriorating with further increases in noise strength.

The CNT-FET signal detector consisted of a chemical vapor deposition (CVD) grown semiconductor carbon nanotube lying on a silicon dioxide insulation layer 500 nanometers (nm) thick and ohmically contacting titanium-gold electrodes (20 nm Ti, 60 nm Au) at both ends (figure S1). The metal contacts were the source and drain electrodes for electric current while the tube was the conduction channel. The p-doped silicon substrate beneath the silicon dioxide layer was the back gate that completed the field effect transistor that was the detector. The single-walled nanotube was three to five micrometers ( $\mu\text{m}$ ) long and less than two nm in diameter according to atomic force microscopy.



**Figure S1 (a):** Detector image. Atomic force microscope (AFM) image of the CNT-FET detector. The detector consisted of a semiconductor carbon nanotube lying across two Ti-Au electrodes (top and bottom). The nanotube was three to five  $\mu\text{m}$  long and less than two nm in diameter according to the AFM. The nanotube was undoped.



**Figure S1 (b):** Nanotube layout. The metal electrodes cover the ends of the carbon nanotube. The nanotube lies on top of the thin (500 nm) silicon oxide layer. The underlying substrate has p-type (holes) doping and can conduct electricity to act as the backgate.

The CNT-FET signal detector was a threshold device and was the nonlinear system in the experiment with approximate threshold voltage  $V_T = -2$  volts. The approximation linearly extrapolated

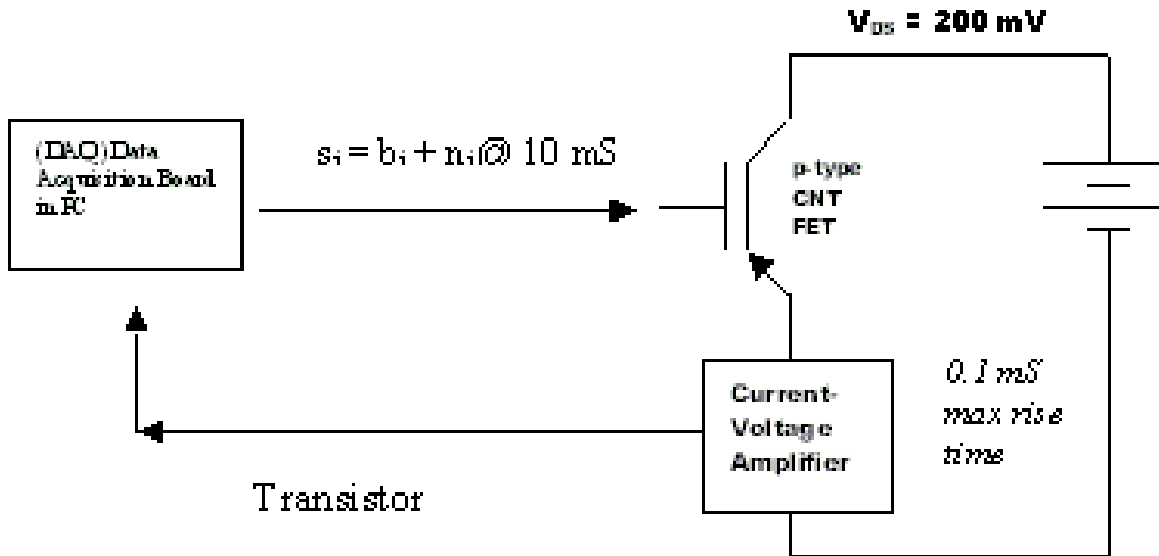
the transistor current-to-gate voltage curve to find a voltage that would intercept the x-axis and would correspond to the OFF state (figure 1 (b) of the main text). The transconductance  $G$  related the output drain-to-source current  $I$  to the input gate voltage  $V$  and the threshold voltage  $V_T$ :

$$I = G (V - V_T) \text{ if } V \leq V_T \text{ and } 0 \text{ otherwise.} \quad (\text{S1})$$

The transconductance  $G$  was negative for the p-type, pristine CNT transistors.

The experiment involved the following equipment (figure S2). A Hewlett Packard 4156 B semiconductor parameter analyzer (not shown) measured  $I-V_D$  and  $I-V_G$  curves that characterized the CNT-FET detector's gate effect. A National Instrument PCI-MI0-16XE-10 multifunction data acquisition (DAQ) board generated the analog voltages that drove the transistor's gate and biased the nanotube then measured the electric current flowing through the nanotube. A DL 1211 current-voltage preamplifier converted the detector's output electric current ( $I_{DS}$ ) to voltage for data acquisition (risetime set to 0.1 ms and sensitivity set to  $10^{-8}$  A/V). Two resistors formed a voltage divider to divide the smallest voltage step by two and improved the resolution of the DAQ's analog voltages. A personal computer running LabView driver controlled the input signal generation and the output measurement to test the CNT-FET detector. A cryostat isolated the detector electrically, kept it at room temperature, and maintained a rough vacuum to remove contaminants such as moisture. A subthreshold gate voltage without additive noise would keep the detector in the OFF state – the drain-to-source current would be in the pico-amp range.





**Figure S2:** Nanotube experiment setup. The threshold detector was a p-type CNT-FET. The input was the gate voltage and the output was the current of the CNT-FET detector. The DAQ board updated the input symbols about once every 10 ms to allow the data acquisition and amplifier hardware to reach steady state. An estimate of each output symbol was the average of 10 measurements that the DAQ made near the end of the symbol interval. Each experiment applied one type of additive white noise for 32 trials 1,000-symbol sequences and used 24 evenly spaced noise values that ranged from 0.001 to 1 standard deviations (dispersions for infinite-variance Cauchy noise).

The experiment generated digital signals in software and converted them to analog voltages to test the detector. An input  $S$  consisted of a sequence of binary symbols  $b$  plus white noise  $n$ :  $s_i = b_i + n_i$ . Each  $b$  was independent, identically distributed (Bernoulli random variables), and took value  $A$  with probability  $p$  or  $\bar{A}$  with probability  $1-p$ . The noise  $n$  was independent of and synchronized with the binary symbols. Each  $n$  was independent and identically distributed. Three types of distributions were

available: Gaussian, uniform, and Cauchy. The binary symbols were subthreshold ( $A = -1.8$  V and  $\bar{A} = -1.6$  V, for example) with respect to the threshold voltage.

An output  $Y$  was the detector's current in response to each input  $S$  at the gate. The p-type transistor model gave  $Y = G(S - V_T)$  if  $S \leq V_T$  and zero otherwise.

The experiment biased the nanotube at 200 mV and updated the input symbol about every 10 ms. The symbol interval was a compromise that produced data in quantity within limited lab time while allowing sufficient time for the preamplifier and DAQ to reach steady state. The experiment conducted 32 trials for each noise type and strength and for each pair of binary symbols. Each trial consisted of a 1000-symbol sequence. The data acquisition equipment measured and averaged ten samples of the detector output near the end of each symbol interval at a rate of 100 kilo-symbols per second to estimate the output symbol sequence. A comparison between the input sequence and the output sequence yielded the system performance.

Cross correlation and mutual information provided comparison between the input and the output sequences and yielded two measures of detector performance. A cross correlation measured the similarity between the input and the output sequences. The correlation measure used the zero-lag value as a scalar representation of the cross correlation sequence between the input and the output:

$$r_{SY}(l) = \sum_{k=1}^N s(k)y(k-l) \quad (\text{S2})$$

$$r_{SY}(0) = \sum_{k=1}^N s(k)y(k) \quad (\text{S3})$$

A normalization scheme divided the zero-lag correlation by the square root of the energy of the input and the output sequences to give the normalized correlation measure where the energy of a sequence is the same as the zero-lag value of its autocorrelation:

$$\begin{aligned}
C(S, Y) &= \frac{\sum_{k=1}^N s(k)y(k)}{\sqrt{\sum_{k=1}^N s(k)s(k)} \sqrt{\sum_{k=1}^N y(k)y(k)}} \\
&= \frac{r_{SY}(0)}{\sqrt{r_{SS}(0)r_{YY}(0)}}
\end{aligned} \tag{S4}$$

This normalized correlation has the maximal value of one for any two identical sequences.

The mutual information was the difference between the output entropy  $H(Y)$  and the conditional output entropy  $H(Y|S)$  conditioned on the input:

$$I(S, Y) = H(Y) - H(Y | S) \tag{S5}$$

where output entropy  $H(Y) = -\sum_{i=1}^N p_i \ln p_i$  used the symbol probabilities  $P(Y = Y_i) = p_i$  and

conditional output entropy  $H(Y | S) = -\sum_{i=1}^N \sum_{j=1}^N p_{ji} \ln \left( \frac{p_{ji}}{p_j} \right)$  used the output probabilities conditioned on

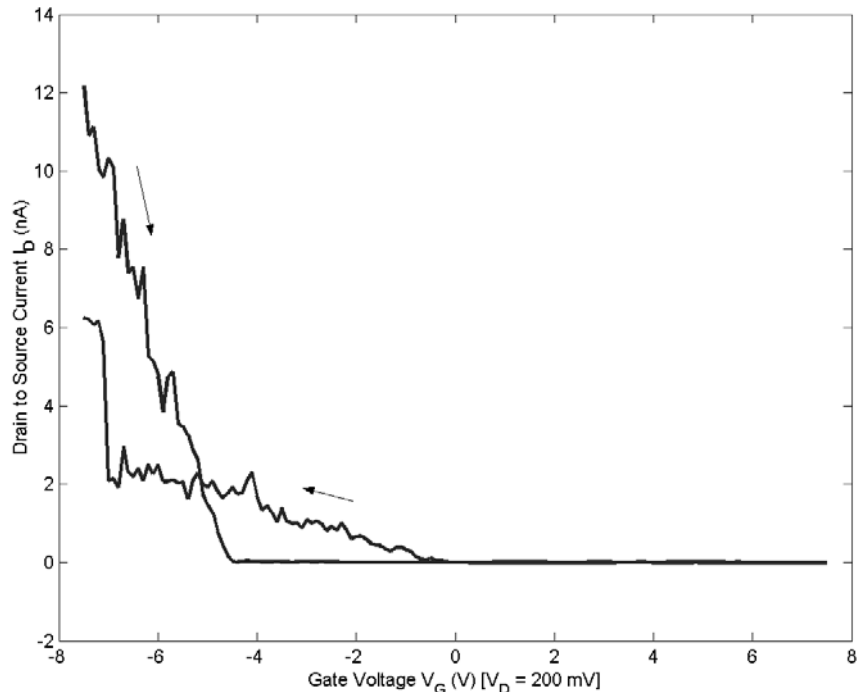
the input  $P_{Y|S}(Y = Y_i | S = S_j) = \frac{p_{ji}}{p_j}$ .

A histogram of each output sequence gave the discrete probability density function to compute the entropies. The use of a histogram avoids imposing an artificial software threshold scheme on the data. The histogram applied a fixed set of bin edges to each output sequence so that the bins represented a fixed set of discrete symbols. The normalization ensured that symbol probability density functions summed to unity.

The nanotube detector exhibited some hysteresis in its gate effect but not enough to prevent the SR effect. The hysteresis (figure S3) affected the transistor's gate effect: threshold voltage shifted based on the direction of the input voltage change. Charge trapping by water molecules on the silicon dioxide surface was one possible mechanism of hysteresis<sup>16</sup>. We kept the detector in vacuum to reduce the

hysteretic effect but some effect persisted even after 72 hours in vacuum. Again the device hysteresis did not prevent the observation of the SR effect.

The experimental data exhibited a gate effect consistent with a transistor in a plot of input sequence versus output sequence (figure 1 (b) of main text). The figure suggested that the signal sequences in the experiment encountered little changes in the threshold effect. The experiment used subthreshold Bernoulli symbols and signals that had short hold times, rapid voltage transitions, and small voltage changes. This voltage scheme differs from the large voltage range and slow voltage-sweep transitions that characterized the hysteresis in figure S3. The experiment yielded evidence of the SR effect in spite of the hysteretic mechanisms.



**Figure S3:** Nanotube transistor gate effect. A semiconductor parameter analyzer produced the  $I-V_G$  curve that showed the transistor current as a function of gate voltage. The gate effect showed some

hysteresis: the threshold voltage and the transconductance varied depending on the direction of the gate voltage sweep. The hysteretic effect did not prevent observation of the SR effect. This set of the  $I-V_G$  curve shows a threshold shift in contrast to the effective  $I-V_G$  curve (figure 1 (b) in main text) that the detector produced in response to the experiment's noisy input signal.

Two statistical tests confirmed that the SR-curves were nonmonotonic. A goodness-of-fit test measures how well a candidate probability density function (pdf) matches a benchmark pdf given a set of data from the candidate pdf. The null hypothesis  $H_0$  states that the two pdfs are the same. The test rejects the null hypothesis if a test statistic exceeds a critical value for a given significance level  $\alpha$ . The significance level  $\alpha$  denotes the probability of a Type-I error--the probability of rejecting the null hypothesis when it is true. The p-value measures the credibility of the null hypothesis  $H_0$  given the data. A statistical test rejects the null hypothesis  $H_0$  at the significance level  $\alpha$  if the p-value is less than the significance level: Reject  $H_0$  if  $p\text{-value} < \alpha$ . Two types of goodness-of-fit tests rejected the match between either SR curve and a monotonically decreasing function based on a  $\beta$ -pdf. The  $\beta$ -pdf (fig. S4 (a)) defines a monotone decreasing function if its two parameters are  $\alpha = 0.5$  and  $\beta = 5$ . It has the form

$$f_{\alpha,\beta}(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} \quad (\text{S6})$$

for  $x \in [0, 1]$  and positive parameters  $\alpha > 0$  and  $\beta > 0$ . The denominator term  $B(\alpha, \beta)$  is

$$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \quad (\text{S7})$$

with  $\Gamma$  function

$$\Gamma(n+1) = \int_0^{\infty} u^n e^{-u} du = n\Gamma(n) \quad (\text{S8})$$

for  $n > 0$  ( $\Gamma(1)=1$  and with  $\Gamma(n+1)=n!$  if  $n$  is a positive integer). The  $\beta$ -pdf contrasts with the SR-curves because it is nonzero only for  $x \in [0, 1]$  and because it decreases monotonically to zero as  $x$  increases to 1 for the parameters  $\alpha = 0.5$  and  $\beta = 5$ .

The goodness-of-fit tests converted each averaged SR-curve to its equivalent pdf  $f_{SR}(k)$ . The conversion interpolated 25 averaged values so that the SR-curves had a uniform increment of  $\Delta x = 0.001$  and were nonzero only in the interval  $[0, 1]$ . The conversion integrated (via discrete approximation) and normalized the SR-curves so that they integrated to one:

$$\int_{-\infty}^{\infty} f_{SR}(x) dx = \int_0^1 f_{SR}(x) dx \approx \sum_{k=1}^N f_{SR}(k) \Delta x = 1 \quad (\text{S9})$$

where  $f_{SR}$  is the normalized SR-curve.

A  $\chi^2$ -test compared the SR-pdfs (mutual information and correlation measure) to the  $\beta$ -pdf in figure S4 (a). We converted the pdf  $f(k)$  to the cumulative distribution function (CDF)  $F(k)$  by integration (via discrete approximation):

$$F(x) = \int_{-\infty}^x f(u) du = \int_0^x f(u) du \approx \sum_{j=1}^k f(j) \Delta x = F(k) \quad (\text{S10})$$

The CDF appeared in both a  $\chi^2$ -test and a Kolmogorov-Smirnov (KS) test: the tests compared the SR-generated CDF to the  $\beta$ -CDF in figure S4 (b).

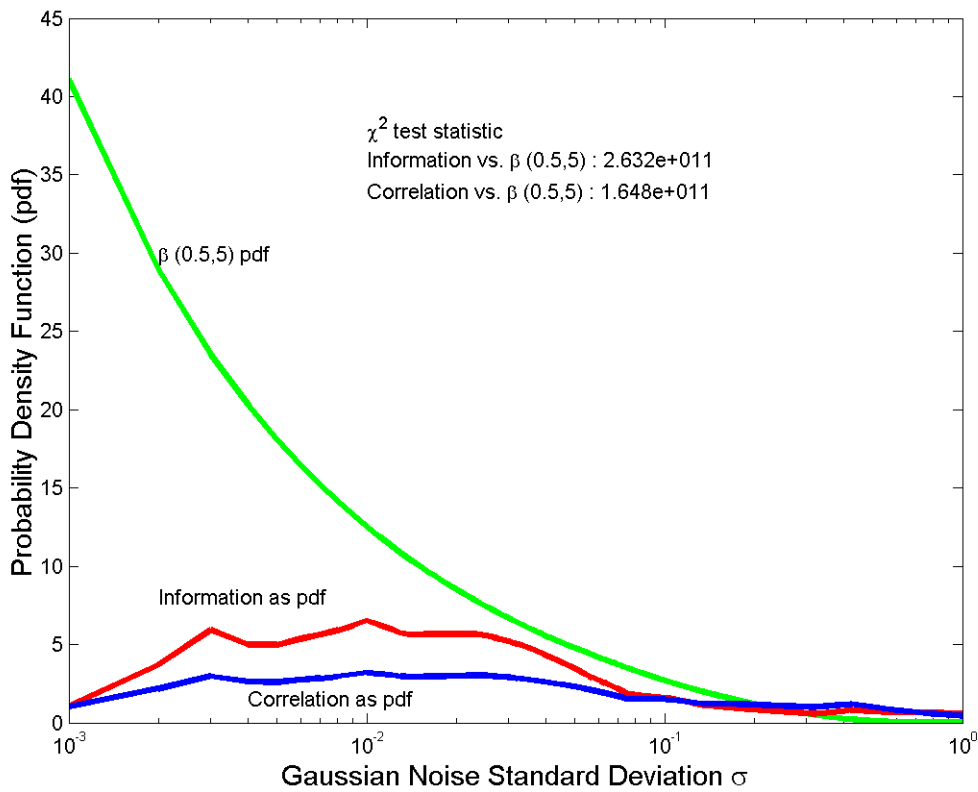


Figure S4 (a) The  $\beta$  probability density function (pdf) and the SR–pdfs. The reference pdf (top green curve) had a  $\beta$  distribution with the parameters  $\alpha = 0.5$  and  $\beta = 5$  ( $\beta(0.5,5)$ ) and decreased monotonically. A normalization scheme converted the SR–curves in figure 1 into their equivalent pdfs  $f_{SR}$  (red for information and blue for correlation). A test statistic that exceeded a critical value rejected the null hypothesis  $H_0$ : SR–pdfs  $\sim \beta(0.5,5)$  and so confirmed that the SR–curves were nonmonotonic.

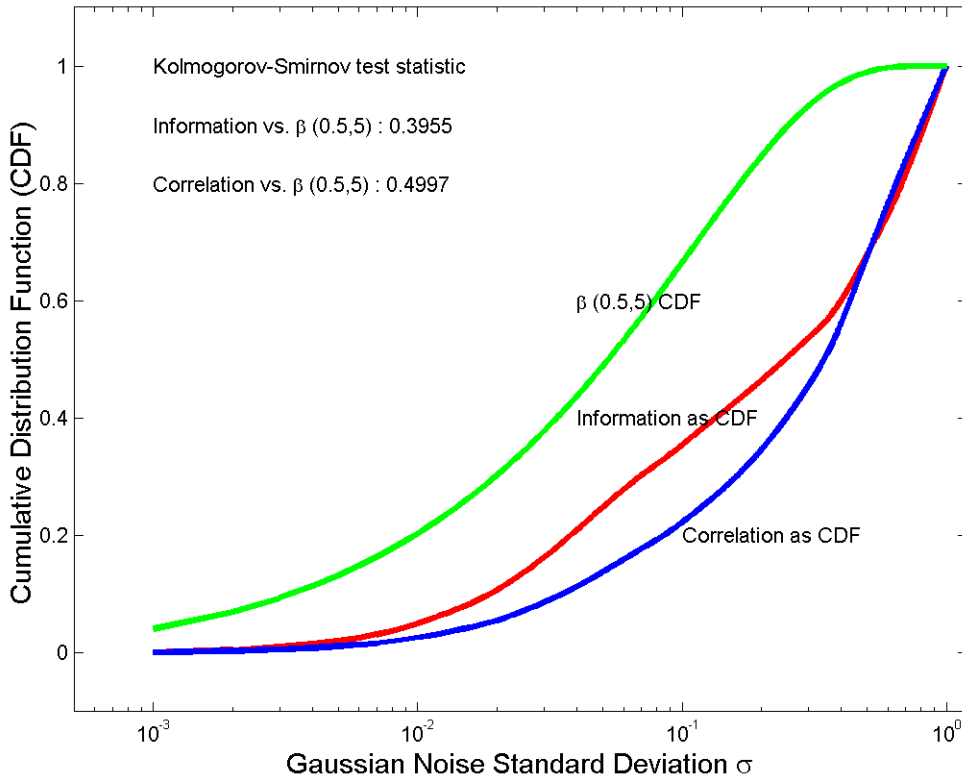


Figure S4 (b) The SR–cumulative distribution functions (CDF) and the  $\beta(0.5,5)$  CDF. Integrating the pdfs gave the CDFs. The  $\beta(0.5,5)$  pdf gave the top green CDF. The SR–pdfs gave the red (information) and blue (correlation) CDFs. The CDF–based statistical tests removed a potential defect in the pdf–based test: the tail of the pdf could skew the test statistic if it gave near–zero values in its denominator. Rejecting the null hypothesis  $H_0$ : SR–CDFs  $\sim \beta(0.5,5)$  confirmed that the SR–curves were nonmonotonic.

The goodness-of-fit test applied a  $\chi^2$ –test with the null hypothesis  $H_0$  : SR–pdfs  $\sim \beta(0.5,5)$  at the smallest level of significance  $\alpha = 0.001$ :

$$H_0 : \text{SR-pdfs} \sim \beta(0.5,5)$$

$$H_a : \text{SR-pdfs} \not\sim \beta(0.5,5)$$



The test rejected the null hypothesis if the test statistic exceeded the critical value. The test statistic had the form

$$\chi_{test}^2 = \sum_i \frac{(O_i - E_i)^2}{E_i} \quad (S11)$$

where  $O_i$  was an observed value in the SR-pdfs and  $E_i$  was an expected value in the reference  $\beta$ -pdf. The critical value was  $\chi_{critical}^2 = 48.2679$  for the smallest level of significance  $\alpha = 0.001$  and for degree of freedom  $\nu = (k - 1 - m) = (25 - 1 - 2) = 22$  where  $k$  was the number of data and  $m$  was the number of parameters in the test. The test statistic was  $\chi_{PDFtest}^2 = 2.632 \times 10^{11}$  for the mutual-information pdf and was  $\chi_{PDFtest}^2 = 1.648 \times 10^{11}$  for the correlation-measure pdf. So the  $\chi^2$ -test showed that the monotonically decreasing  $\beta$ -pdf differed substantially from both the SR-pdfs with  $p$ -value  $p < 0.001$ .

A second  $\chi^2$ -test based on the CDF removed a potential confounding factor in the pdf-based test: the small values in the tail of the pdf might skew the test statistic if it gave near-zero values in its denominator. The CDF-based goodness-of-fit test applied the same hypotheses as the pdf-based test but changed the test statistic to use the observed and expected CDF values:

$$\begin{aligned} \mathbf{H}_0 &: \text{SR-CDFs} \sim \beta(0.5,5) \\ \mathbf{H}_a &: \text{SR-CDFs} \not\sim \beta(0.5,5) \end{aligned}$$

instead of the pdf. The test statistic was  $\chi_{CDFtest}^2 = 89.2559$  for the mutual-information CDF and was  $\chi_{CDFtest}^2 = 129.1207$  for the correlation-measure CDF. Both test statistics greatly exceeded the critical value  $\chi_{critical}^2 = 48.2679$ . So the  $\chi^2$ -test also showed that the  $\beta$ -CDF differed substantially from both SR-CDFs with  $p$ -value  $< 0.001$ .

The Kolmogorov-Smirnov (KS) test for goodness-of-fit also tested how well the SR-CDFs matched a  $\beta$ -CDF:

$$\mathbf{H}_0 : \text{SR-CDFs} \sim \beta(0.5,5)$$

$$\mathbf{H}_a : \text{SR-CDFs} \not\sim \beta(0.5,5)$$

by comparing the CDF-based test statistic to the critical value  $KS_{critical} = 0.32$  for the smallest significance level  $\alpha = 0.01$  and for  $n = 25$  (number of data). The test statistic equaled the largest difference between the observed and the expected CDF values:

$$KS_{test} = \max_i (|O_i - E_i|) \quad (\text{S12})$$

where  $O_i$  was an observed value in the SR-CDF and  $E_i$  was an expected value in the reference  $\beta$ -CDF. Both test statistics exceeded the critical value:  $KS_{test} = 0.3955$  for the mutual-information CDF and  $KS_{test} = 0.4997$  for the correlation-measure CDF. So the KS-test rejected the null hypothesis and showed that the monotonic decreasing  $\beta$ -CDF differed from both the SR-CDFs with  $p$ -value  $< 0.01$ .