Fisher Information as a Metric of Locally Optimal Processing and Stochastic Resonance

Fabing Duan[1]*, **François Chapeau-Blondeau**[2], **Derek Abbott**[3]

1 College of Automation Engineering, Qingdao University, Qingdao, People's Republic of China, **2** Laboratoire d'Ingénierie des Systèmes Automatisés (LISA), Université d'Angers, Angers, France, **3** Centre for Biomedical Engineering (CBME) and School of Electrical & Electronic Engineering, The University of Adelaide, Adelaide, Southern Australia, Australia

Abstract

The origins of Fisher information are in its use as a performance measure for parametric estimation. We augment this and show that the Fisher information can characterize the performance in several other significant signal processing operations. For processing of a weak signal in additive white noise, we demonstrate that the Fisher information determines (i) the maximum output signal-to-noise ratio for a periodic signal; (ii) the optimum asymptotic efficacy for signal detection; (iii) the best cross-correlation coefficient for signal transmission; and (iv) the minimum mean square error of an unbiased estimator. This unifying picture, via inequalities on the Fisher information, is used to establish conditions where improvement by noise through stochastic resonance is feasible or not.

Citation: Duan F, Chapeau-Blondeau F, Abbott D (2012) Fisher Information as a Metric of Locally Optimal Processing and Stochastic Resonance. PLoS ONE 7(4): e34282. doi:10.1371/journal.pone.0034282

Editor: Matjaz Perc, University of Maribor, Slovenia

Received February 1, 2012; **Accepted** February 25, 2012; **Published** April 6, 2012

Funding: This work is sponsored by the NSF of Shandong Province, China (No. ZR2010FM006). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: fabing.duan@gmail.com

Introduction

Fisher information is foremost a measure of the minimum error in estimating an unknown parameter of a probability distribution, and its importance is related to the Cramér-Rao inequality for unbiased estimators [1,2]. By introducing a location parameter, the de Bruijn's identity indicates that the fundamental quantity of Fisher information is affiliated with the differential entropy of the minimum descriptive complexity of a random variable [1]. Furthermore, in known weak signal detection, a locally optimal detector, acting as the small-signal limited Neyman-Pearson detector, has favorable properties for small signal-to-noise ratios [3]. With sufficiently large observed data and using the central limit theorem, it is demonstrated that the locally optimal detector is asymptotically optimum and the Fisher information of the noise distribution is the upper bound of the asymptotic efficacy [2–7]. For weak random signal detection, the second order Fisher information is also associated with the maximum asymptotic efficacy of the generalized energy detector [4–7].

However, the fundamental nature of Fisher information is not adequately recognized for processing weak signals. To extend the heuristic studies of [1–7], in this paper, we will theoretically demonstrate that, for a weak signal buried in additive white noise, the performance for locally optimal processing can be generally measured by the Fisher information of the noise distribution. We show this for the following signal processing case studies: (i) the maximum output signal-to-noise ratio for a periodic signal; (ii) the optimum asymptotic efficacy for signal detection; (iii) the best cross-correlation coefficient for signal transmission; and (iv) the minimum mean square error of an unbiased estimator. The physical significance of Fisher information is that it provides a unified bound for characterizing the performance for locally optimal processing. Furthermore, we establish the Fisher information condition for stochastic resonance (SR) that has been studied for improving system performance over several decades [8–32]. In our recent work [28], it is established that improvement by adding noise is impossible for detecting a weak known signal. Here, based on Fisher information inequalities, we further prove that SR is not applicable for improving the performance of locally optimal processing in the considered cases (i)–(iv). This result generalizes a proof that existed previously only for a weak periodic signal in additive Gaussian noise [12,33]. However, beyond these restrictive conditions, the observed noise-enhanced effects [9–11,26,28–30] show that SR can provide a signal processing enhancement using the constructive role of noise. The applications of SR to nonlinear signal processing are of practical interest.

Results

In many situations we are interested in processing signals that are very weak compared to the noise level [2,3,6]. It would be desirable in these situations to determine an optimal memoryless nonlinearity in the following study cases.

Output signal-to-noise ratio for a periodic signal

First, consider a static nonlinearity with its output

$$y(t) = g[x(t)], \tag{1}$$

where the function g is a memoryless nonlinearity and the input is a signal-plus-noise mixture $x(t) = s(t) + z(t)$. The component $s(t)$ is a known weak periodic signal with a maximal amplitude A

$(0 \leq |s(t)| \leq A)$ and period T. Zero-mean white noise $z(t)$, independent of $s(t)$, has probability density function (PDF) f_z and a root-mean-square (RMS) amplitude σ_z. It is assumed that g has zero mean under f_z, i.e. $\int_{-\infty}^{\infty} g(x) f_z(x) dx = \mathrm{E}[g(x)] = 0$, which is not restrictive since any arbitrary g can always include a constant bias to cancel this average [6]. The input signal-to-noise ratio for $x(t)$ can be defined as the power contained in the spectral line $1/T$ divided by the power contained in the noise background in a small frequency bin ΔB around $1/T$ [10], this is

$$R_{\text{in}} = \frac{|\langle s(t) \exp[-i2\pi t/T] \rangle|^2}{\sigma_z^2 \Delta B \Delta t}, \tag{2}$$

with Δt indicating the time resolution or the sampling time in a discrete-time implementation and the temporal average defined as $\langle \cdots \rangle = \frac{1}{T}\int_0^T \cdots dt$ [10]. Here, we assume the sampling time $\Delta t \ll T$ and observe the output $y(t)$ for a sufficiently large time interval of NT $(N \gg 1)$ [10]. Since $s(t)$ is periodic, $y(t)$ is in general a cyclostationary random signal with period T [10]. Similarly, the output signal-to-noise ratio for $y(t)$ is given by

$$R_{\text{out}} = \frac{|\langle \mathrm{E}[y(t)] \exp[-i2\pi t/T] \rangle|^2}{\langle var[y(t)] \rangle \Delta B \Delta t}, \tag{3}$$

with nonstationary expectation $\mathrm{E}[y(t)]$ and nonstationary variance $\mathrm{var}[y(t)]$ [10].

In the case of $A \to 0$, we have a Taylor expansion of the expectation at a fixed time t as

$$\begin{aligned} \mathrm{E}[y(t)] &= \int_{-\infty}^{\infty} g(x) f_z(x-s) dx \approx \int_{-\infty}^{\infty} g(x)[f_z(x) - s(t) f'_z(x)] dx \\ &= s(t) \int_{-\infty}^{\infty} g'(x) f_z(x) dx = s(t) \mathrm{E}[g'(x)], \end{aligned} \tag{4}$$

where we assume the derivatives $g'(x) = dg(x)/dx$ and $f'_z(x) = df_z(x)/dx$ exist for almost all x (similarly hereinafter) [2,6]. Thus, we have

$$\begin{aligned} \mathrm{var}[y(t)] &= \mathrm{E}[y^2(t)] - \mathrm{E}[y(t)]^2 \approx \mathrm{E}[y^2(t)] - s^2(t)\mathrm{E}^2[g'(x)] \\ &\approx \int_{-\infty}^{\infty} g^2(x)[f_z(x) - s(t) f'_z(x)] dx \\ &\approx \int_{-\infty}^{\infty} g^2(x) f_z(x) dx = \mathrm{E}[g^2(x)], \end{aligned} \tag{5}$$

where $s(t)\int_{-\infty}^{\infty} g^2(x) f'_z(x) dx = 2s(t)\mathrm{E}[g(x)g'(x)]$ and $s^2(t)\mathrm{E}^2[g'(x)]$, compared with $\mathrm{E}[g^2(x)]$, can be neglected as $A \to 0$ $(0 < |s(t)| \leq A)$ [2,6]. The above derivations of Eqs. (4) and (5) are exact in the asymptotic limit for weak signals, and have been generally adopted in [2,6].

Substituting Eqs. (4) and (5) into Eq. (3), we have

$$\begin{aligned} R_{\text{out}} &\approx \frac{|\langle s(t) \exp[-i2\pi t/T] \rangle|^2}{\Delta B \Delta t} \frac{\mathrm{E}^2[g'(x)]}{\mathrm{E}[g^2(x)]} \\ &\leq \frac{|\langle s(t) \exp[-i2\pi t/T] \rangle|^2}{\Delta B \Delta t} \mathrm{E}\left[\frac{f'^2_z(x)}{f_z^2(x)}\right] \\ &= \frac{|\langle s(t) \exp[-i2\pi t/T] \rangle|^2}{\Delta B \Delta t} I(f_z), \end{aligned} \tag{6}$$

where the expectation $\mathrm{E}\left[f'^2_z(x)/f_z^2(x)\right]$ is simply the Fisher information $I(f_z)$ of the noise PDF f_z [2,6], and the equality occurs as

$$g(x) = C f'_z(x)/f_z(x) \triangleq g_{\text{opt}}(x), \tag{7}$$

by the Cauchy-Schwarz inequality for a constant C [2,6].

Noting Eqs. (2) and (6), the output-input signal-to-noise ratio gain G is bounded by

$$G = \frac{R_{\text{out}}}{R_{\text{in}}} \approx \sigma_z^2 \frac{\mathrm{E}^2[g'(x)]}{\mathrm{E}[g^2(x)]} \leq \sigma_z^2 \mathrm{E}\left[\frac{f'^2_z(x)}{f_z^2(x)}\right] = \sigma_z^2 I(f_z) = I(f_{z_0}), \tag{8}$$

with equality achieved when g takes the locally optimal nonlinearity g_{opt} of Eq. (7). Here, for a standardized PDF f_{z_0} with zero mean and unity variance $\sigma_{z_0}^2 = 1$, the scaled noise $z(t) = \sigma_z z_0(t)$ has its PDF $f_z(z) = f_{z_0}(z/\sigma_z)/\sigma_z$ and the Fisher information satisfies $I(f_z) = I(f_{z_0})/\sigma_z^2$ [1,34]. It is known that a standardized Gaussian PDF $f_{z_0}(z_0) = \exp(-z_0^2/2)/\sqrt{2\pi}$ has the minimal Fisher information $I(f_{z_0}) = 1$ and any standardized non-Gaussian PDF f_{z_0} has the Fisher information $I(f_{z_0}) > 1$ [2]. It can be seen that, the linear system $g_{\mathrm{L}}(x) = x$ has its output signal-to-noise ratio $R_{\text{out}} = R_{\text{in}}$ in Eq. (3). Thus, the output-input signal-to-noise ratio gain G in Eq. (8) also clearly represents the expected performance improvement of the nonlinearity g over the linear system g_{L}.

Optimum asymptotic efficacy for signal detection

Secondly, we consider the observation vector $X = (X_1, X_2, \cdots, X_N)$ of real-valued components X_n by

$$X_n = \theta s_n + z_n, \quad n = 1, 2, \cdots, N, \tag{9}$$

where the components z_n form a sequence of independent and identically distributed (i.i.d.) random variables with PDF f_z, and the known signal components s_n are with the signal strength θ [6]. For the known signal sequence $\{s_n, n = 1, 2, \cdots, N\}$, it is assumed that there exists a finite (non-zero) bound A such that $0 \leq |s_n| \leq A$, and the asymptotic average signal power is finite and non-zero, i.e. $0 < P_s^2 = \lim_{N\to\infty} \sum_{n=1}^N s_n^2/N < \infty$ [6]. Then, the detection problem can be formulated as a hypothesis-testing problem of deciding a null hypothesis H_0 $(\theta = 0)$ and an alternative hypothesis H_1 $(\theta > 0)$ describing the joint density function of X with

$$\begin{aligned} H_0 &: f_X(X) = \prod_{n=1}^N f_z(X_n) \ \text{ for } \ \theta = 0; \\ H_1 &: f_X(X) = \prod_{n=1}^N f_z(X_n - \theta s_n) \ \text{ for } \ \theta > 0. \end{aligned} \tag{10}$$

Consider a generalized correlation detector

$$T_{\mathrm{GC}}(X) = \sum_{n=1}^N g(X_n) s_n \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma, \tag{11}$$

where the memoryless nonlinearity g has zero mean under f_z, i.e. $\mathrm{E}[g(x)] = 0$ [6]. In the asymptotic case of $\theta \to 0$ and $N \to \infty$, the test statistic T_{GC}, according to the central limit theorem, converges to a Gaussian distribution with mean $\mathrm{E}[T_{\mathrm{GC}}|H_0] = 0$ and variance $\mathrm{var}[T_{\mathrm{GC}}|H_0] \approx N P_s^2 \mathrm{E}[g^2(x)]$ under the null hypotheses H_0 [6]. Using Eqs. (4) and (5), T_{GC} is asymptotically Gaussian with mean $\mathrm{E}[T_{\mathrm{GC}}|H_1] \approx \theta N P_s^2 \mathrm{E}[g'(x)]$ and variance $\mathrm{var}[T_{\mathrm{GC}}|H_1] = \mathrm{var}[T_{\mathrm{GC}}|H_0]$ under the hypothesis H_1 [6].

Given a false alarm probability P_{FA}, the asymptotic detection probability P_{D} for the generalized correlation detector of Eq. (11) can be expressed as [2,6]

$$P_{\mathrm{D}} = Q[Q^{-1}(P_{\mathrm{FA}}) - \sqrt{N}\theta P_s \sqrt{\xi_{\mathrm{GC}}}], \quad (12)$$

with $Q(x) = \int_x^\infty \exp[-t^2/2]/\sqrt{2\pi}\, dt$ and its inverse function Q^{-1} [2,6]. Thus, for fixed N and θP_s (since the signal is known), P_{D} is a monotonically increasing function of the normalized asymptotic efficacy ξ_{GC} given by [6]

$$\xi_{\mathrm{GC}} = \lim_{N\to\infty} \frac{\{\frac{d\mathrm{E}[T_{\mathrm{GC}}(X)]}{d\theta}|_{\theta=0}\}^2}{P_s^2 N \mathrm{var}[T_{\mathrm{GC}}(X)]|_{\theta=0}} = \frac{\mathrm{E}^2[g'(x)]}{\mathrm{E}[g^2(x)]} \le \mathrm{E}\left[\frac{f'^2_z(x)}{f_z^2(x)}\right] = I(f_z), \quad (13)$$

with equality being achieved when $g = g_{\mathrm{opt}}$ in Eq. (7). This result also indicates that the asymptotic optimal detector is just the locally optimal detector established by the Taylor expansion of the likelihood ratio test statistic $\ln[\Pi_{n=1}^N f_z(X_n - \theta s_n)/\Pi_{n=1}^N f_z(X_n)] \approx \sum_{n=1}^N g_{\mathrm{opt}}(X_n)\theta s_n$ $(C=-1)$ in terms of the generalized Neyman-Pearson lemma [2,6].

Interestingly, with $\xi_{\mathrm{LC}} = \mathrm{E}^2[g'(x)]/\mathrm{E}[g^2(x)] = \sigma_z^{-2}$ achieved by a linear correlation detector ($g_{\mathrm{LC}}(x) = x$ in Eq. (11)) as a benchmark [5,6], the asymptotic relative efficiency

$$\mathrm{ARE} = \frac{\xi_{\mathrm{GC}}}{\xi_{\mathrm{LC}}} = \sigma_z^2 \frac{\mathrm{E}^2[g'(x)]}{\mathrm{E}[g^2(x)]} \le \sigma_z^2 I(f_z) = I(f_{z_0}), \quad (14)$$

provides an asymptotic performance improvement of a generalized correlation detector over the linear correlation detector when both detectors operate in the same noise environment [5,6].

Next, consider the weak random signal components s_n has PDF f_s with zero mean $\int_{-\infty}^\infty s_n f_s ds = 0$ and variance $\sigma_s^2 = \int_{-\infty}^\infty s_n^2 f_s ds = 1$ in the observation model of Eq. (9) [5,6]. Here, the signal components s_n are i.i.d. Then, this random signal hypothesis test becomes [6]

$$H_0 : f_X(X) = \prod_{n=1}^N f_z(X_n), \text{ for } \theta = 0;$$
$$H_1 : f_X(X) = \int_{-\infty}^\infty \prod_{n=1}^N f_z(X_n - \theta s_n) f_s(s_n) ds_n, \text{ for } \theta > 0,$$

for determining whether the random signal is present or not. Consider a generalized energy detector

$$T_{\mathrm{GE}}(X) = \sum_{n=1}^N g(X_n) \underset{H_0}{\overset{H_1}{\gtrless}} \gamma, \quad (15)$$

where we also assume $\mathrm{E}[T_{\mathrm{GE}}|H_0] = 0$, and then $\mathrm{var}[T_{\mathrm{GE}}|H_0] = N\mathrm{E}[g^2(x)]$. Furthermore, in the asymptotic case of $\theta \to 0$, the expectation [6]

$$\mathrm{E}[T_{\mathrm{GE}}|H_1] = N\int_{-\infty}^\infty g(x) \int_{-\infty}^\infty f_z(x-\theta s) f_s(s) ds\, dx$$
$$\approx N \int_{-\infty}^\infty g(x) \int_{-\infty}^\infty [f_z(x) - \theta s f'_z(x) + \frac{\theta^2 s^2}{2} f''_z(x)] f_s(s) ds\, dx \quad (16)$$
$$\approx \frac{N\theta^2}{2} \int_{-\infty}^\infty g(x) f''_z(x) dx = \frac{N\theta^2}{2} \mathrm{E}[g''(x)].$$

Thus, the efficacy of a generalized energy detector is defined as [6]

$$\xi_{\mathrm{GE}} = \lim_{N\to\infty} \frac{\{\frac{d\mathrm{E}[T_{\mathrm{GE}}(X)]}{d\theta^2}|_{\theta^2=0}\}^2}{N \mathrm{var}[T_{\mathrm{GE}}(X)]|_{\theta^2=0}} = \frac{1}{4}\frac{\mathrm{E}^2[g''(x)]}{\mathrm{E}[g^2(x)]} \le \frac{1}{4}\mathrm{E}\left[\frac{f''^2_z(x)}{f_z^2(x)}\right] = \frac{1}{4} I_2(f_z), \quad (17)$$

where θ^2 is treated as the signal strength parameter and $I_2(f_z)$ is the second order Fisher information [6,7]. It is noted that the equality of Eq. (17) is achieved as $g(x) = g_{\mathrm{opt}}(x) = Cf''_z/f_z$ for a constant C [6]. Given a false alarm probability P_{FA}, the asymptotic detection probability P_{D} for the generalized energy detector of Eq. (15) is a monotonically increasing function of the efficacy ξ_{GE} [5–7].

Cross-correlation coefficient for signal transmission

Thirdly, we transmit a weak aperiodic signal $s(t)$ through the nonlinearity g of Eq. (1) [13]. Here, the signal $s(t)$ is with the average signal variance $\sigma_s^2 \ll \sigma_z^2$, the zero mean and the upper bound A $(0 \le |s(t)| \le A)$. For example, $s(t)$ can be a sample according to a uniformly distributed random signal equally taking values from a bounded interval. The input cross-correlation coefficient of $s(t)$ and $x(t) = s(t) + z(t)$ is defined as [2,13]

$$\rho_{s,x} = \frac{\mathrm{E}[s(t)x(t)]}{\sigma_s \sqrt{\mathrm{E}[x^2(t)]}} = \frac{\frac{\sigma_s}{\sigma_z}}{\sqrt{\frac{\sigma_s^2}{\sigma_z^2} + 1}} \approx \frac{\sigma_s}{\sigma_z}. \quad (18)$$

Using Eqs. (4) and (5), the output cross-correlation coefficient of $s(t)$ and $y(t) = g[x(t)]$ is given by

$$\rho_{s,y} = \frac{\mathrm{E}[s(t)y(t)]}{\sigma_s \sqrt{\mathrm{var}[y(t)]}} \approx \frac{\sigma_s \mathrm{E}[g'(x)]}{\sqrt{\mathrm{E}[g^2(x)]}} \le \sigma_s \sqrt{I(f_z)}, \quad (19)$$

which has its maximal value as $g = g_{\mathrm{opt}}$ of Eq. (7). Then, the cross-correlation gain G_ρ is bounded by

$$G_\rho = \frac{\rho_{s,y}}{\rho_{s,x}} \approx \sigma_z \frac{\mathrm{E}[g'(x)]}{\sqrt{\mathrm{E}[g^2(x)]}} \le \sqrt{I(f_{z_0})}. \quad (20)$$

Mean square error of an unbiased estimator

Finally, for the N observation components $x_n = s_n(\theta) + z_n$, we assume the signal $s_n(\theta)$ are with an unknown parameter θ. As the upper bound $A \to 0$ $(0 \le |s_n| \le A)$, the Cramér-Rao inequality indicates that the mean squared error of any unbiased estimator of the parameter θ is lower bounded by the reciprocal of the Fisher information [1,2] given by

$$\begin{aligned} I(\theta) &= \sum_{n=1}^N \mathrm{E}\left[\left(\frac{\partial \ln f_z(x_n - s_n)}{\partial\theta}\right)^2\right] \\ &\approx \sum_{n=1}^N \mathrm{E}\left[\left(\frac{df_z(z_n)/dz_n}{f_z(z_n)}|_{z_n = x_n - s_n}(-\frac{\partial s_n}{\partial\theta})\right)^2\right] \quad (21) \\ &= I(f_z)\sum_{n=1}^N (\frac{\partial s_n}{\partial\theta})^2, \end{aligned}$$

which indicates that the minimum mean square error of any unbiased estimator is also determined by the Fisher information $I(f_z)$ of a distribution, as $\sum_{n=1}^N (\frac{\partial s_n}{\partial\theta})^2$ is given.

Therefore, just as the Fisher information represents the lower bound of the mean squared error of any unbiased estimator in signal estimation [1,2], the physical significance of the Fisher information $I(f_z)(I_2(f_z))$ is that it provides a unified upper bound of the performance for locally optimal processing in the considered signal processing cases.

Aiming to explain the upper bound of the performance for locally optimal processing as Fisher information, we here show an illustrative example in Fig. 1. Consider the generalized Gaussian noise with PDF

$$f_z(x)=\frac{c_1}{\sigma_z}\exp\left(-c_2\left|\frac{x}{\sigma_z}\right|^{\alpha}\right), \qquad (22)$$

where $c_1=\frac{\alpha}{2}\Gamma^{\frac{1}{2}}(3\alpha^{-1})/\Gamma^{\frac{3}{2}}(\alpha^{-1})$, $c_2=\left[\Gamma(3\alpha^{-1})/\Gamma(\alpha^{-1})\right]^{\frac{\alpha}{2}}$ for a rate of exponential decay parameter $\alpha>0$ [2,6]. The corresponding locally optimal nonlinearity is $g_{opt}(x)=|x|^{\alpha-1}sign(x)$ and the output-input signal-to-noise ratio gain in Eq. (8) is $G=I(f_{z_0})=\alpha^2\Gamma(3\alpha^{-1})\Gamma(2-\alpha^{-1})/\Gamma^2(\alpha^{-1})$ (solid line), as shown in Fig. 1. For comparison, we also operate the sign nonlinearity $g_{\mathrm{S}}(x)=\mathrm{sign}(x)$ and the linear system $g_{\mathrm{L}}(x)=x$ in the generalized Gaussian noise. The output-input signal-to-noise ratio gain in Eq. (8) of g_{S} is $G=4\sigma_z^2 f_z^2(0)=4f_{z_0}^2(0)$ (dashed line), as shown in Fig. 1. For the linear system g_{L}, Eq. (8) indicates that $G=1$ (dotted line) for $\alpha>0$, as plotted in Fig. 1. It is seen in Fig. 1 that, only for $\alpha=1$, the performance of g_{S} attains that of the locally optimal nonlinearity of g_{opt}. This is because, the nonlinearity g_{S} is just the locally optimal nonlinearity for Laplacian noise ($\alpha=1$), and the Fisher information limit $I(f_{z_0})=2$ is achieved. Likewise, for Gaussian noise ($\alpha=2$), the linear system g_{L} is optimal and the output-input SNR gain $G=I(f_{z_0})=1$. It is noted that the above analyses are also valid for the asymptotic relative efficiency of Eq. (14) and the cross-correlation gain of Eq. (20).

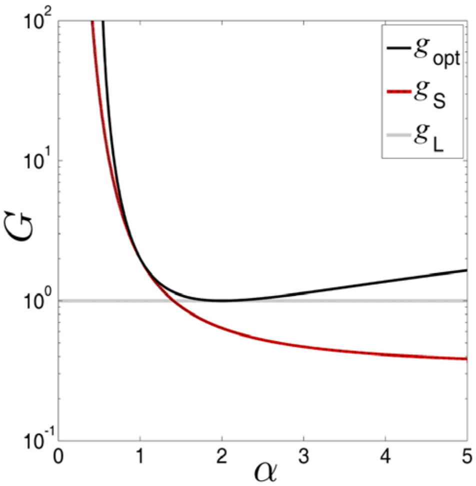

Figure 1. The output-input signal-to-noise ratio gain G. The output-input signal-to-noise ratio gain G versus the exponential decay parameter α of the generalized Gaussian noise for the locally optimal nonlinearity g_{opt} (solid line), the sign nonlinearity g_{S} (red line) and the linear system g_{L} (dotted line), respectively.
doi:10.1371/journal.pone.0034282.g001

Fisher information condition for stochastic resonance

Stochastic resonance (SR), being contrary to conventional approaches of suppressing noise, adds an appropriate amount of noise to a nonlinear system to improve its performance [8–32]. SR emerged from the field of climate dynamics [8], and the topic has flourished in physics [15–19] and neuroscience [13,14,20]. The notion of SR has been widened to include a number of different mechanisms [15,17,25], and SR effects have also been demonstrated in various extended systems [9–20,25] and complex networks [21–24,27].

An open question concerning SR is that, under the asymptotic cases of weak signal and large sample size, can SR play a role in locally optimal processing? Here, based on the Fisher information inequalities, we will demonstrate that SR is inapplicable to performance improvement for locally optimal processing.

For a given observation $x(t)=s(t)+z(t)$, we add the extra noise $v(t)$, independent of the initial noise $z(t)$ and the signal $s(t)$, to $x(t)$. Then, the updated data $\hat{x}(t)=s(t)+z(t)+v(t)=s(t)+w(t)$. Here, the composite noise $w(t)$ has a convolved PDF

$$f_w(x)=\int_{-\infty}^{\infty} f_z(x-u)f_v(u)du, \qquad (23)$$

where f_v is the PDF of noise $v(t)$. Currently, the weak signal $s(t)$ is corrupted by the composite noise $w(t)$, and then the performance measures of locally optimal processing in Eqs. (6), (13), (17), (19) and (21) should be replaced with $I(f_w)$ $(I_2(f_w))$. It can be shown by the Cauchy-Schwarz inequality that [34]

$$I(f_w)\leq \min(I(f_z),I(f_v)), \qquad (24)$$

$$I_2(f_w)\leq \min(I_2(f_z),I_2(f_v)). \qquad (25)$$

This is because that, if $I(f_z)\leq I(f_v)$, then using $f'_w(x)=\int_{-\infty}^{\infty} f'_z(x-u)f_v(u)du$ and the Cauchy-Schwarz inequality [34]

$$\begin{aligned} I(f_w) &= \int \frac{(f'_w(x))^2}{f_w(x)}dx \\ &= \int\left\{\frac{\left[\int (f'_z(x-u)/f_z(x-u))f_z(x-u)f_v(u)du\right]^2}{f_w(x)}\right\}dx \\ &\leq \int\left\{\int\left(\frac{f'_z(x-u)}{f_z(x-u)}\right)^2 f_z(x-u)f_v(u)du\right\}dx \\ &= \int\int\left(\frac{f'_z(z)}{f_z(z)}\right)^2 f_z(z)f_v(u)dudz=I(f_z). \end{aligned} \qquad (26)$$

Similarly, substituting $f''_w(x)=\int_{-\infty}^{\infty} f''_z(x-u)f_v(u)du$ into Eq. (26), we also obtain $I_2(f_w)\leq I_2(f_z)$ of Eq. (25).

Therefore, in asymptotic cases of weak signal and large sample size, Eqs. (24) and (25) show that SR cannot improve the performance of the above four locally optimal processing cases by adding more noise. However, the asymptotic limits of weak signal and large sample size are well delimited, and may not be met in practice. It is interesting to note that, under less restrictive conditions, noise-enhanced effects have been observed in fixed locally optimal detectors [9], suboptimal detectors [26,29], the optimal detector with finite sample sizes [11] or non-weak signals [11,25], soft-threshold systems [30] and the dead-zone limiter detector [28] by utilizing the constructive role of noise.

We here present an illustrative example of SR that occurs outsides restrictive conditions, where a suboptimal detector is adopted for Gaussian noise. Consider a generalized correlation

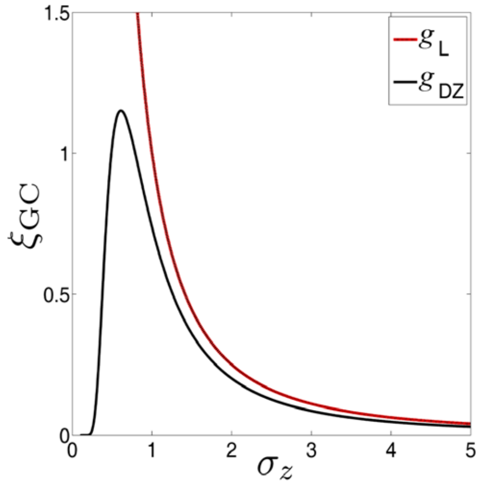

Figure 2. The normalized asymptotic efficacy ξ_{GC}. The normalized asymptotic efficacy ξ_{GC} of the dead-zone limiter nonlinearity g_{DZ} (solid line) and the linear system g_L (red line) as a function of the RMS amplitude σ_z of Gaussian noise ($\alpha = 2$).
doi:10.1371/journal.pone.0034282.g002

detector of Eq. (11) based on the dead-zone limiter nonlinearity

$$g_{DZ}(x) = \begin{cases} -1 & \text{for } x < -c, \\ 0 & \text{for } -c \le x \le c, \\ +1 & \text{for } x > c, \end{cases} \tag{27}$$

with response thresholds at $x = \pm c$ [6]. For the generalized Gaussian noise of Eq. (22), the normalized asymptotic efficacy ξ_{GC} in Eq. (13) of g_{DZ} can be rewritten as

$$\xi_{GC} = \frac{1}{c^2} (\frac{c}{\sigma_z})^2 \frac{2 f_{z_0}^2 (c/\sigma_z)}{1 - F_{z_0}(c/\sigma_z)}, \tag{28}$$

where F_{z_0} is the cumulative distribution function of the standardized generalized Gaussian noise PDF f_{z_0} [28]. For a fixed response threshold c ($c = 1$ without loss of generality), we plot the the normalized asymptotic efficacy ξ_{GC} (solid line) of the dead-zone limiter nonlinearity g_{DZ} as a function of the RMS amplitude σ_z of Gaussian noise ($\alpha = 2$), as shown in Fig. 2. It is clearly seen in Fig. 2 that the SR effect appears, and ξ_{GC} achieves its maximum $\xi_{GC}^* = 1.1512$ at a non-zero level of $\sigma_z^* = 0.6098$. If the original Gaussian noise RMS $\sigma_z < \sigma_z^* = 0.6098$, we can add independent Gaussian noise $v(t)$ with its RMS amplitude $\sigma_v = \sqrt{\sigma_z^{*2} - \sigma_z^2}$ to increase ξ_{GC} to the maximum $\xi_{GC}^* = 1.1512$ [28]. However, g_{DZ} is a suboptimal nonlinearity for Gaussian noise, and the locally optimal detector is the linear correlation detector based on the linear system $g_L(x) = x$ in Eq. (11). It is seen in Fig. 2 that g_{DZ} can not overperform g_L (dashed line), even we can add the appropriate amount of noise to exploit constructive role of noise in g_{DZ}.

Discussion

In this paper, for a weak signal in additive white noise, it is theoretically demonstrated that the optimum performance for locally optimal processing is upper bounded by the Fisher information of the noise distribution, and this is uniformly obtained in (i) the maximum output signal-to-noise ratio ratio for a periodic signal; (ii) the optimum asymptotic efficacy for signal detection; (iii) the best cross-correlation coefficient for signal transmission; and (iv) the minimum mean square error of an unbiased estimator. Based on the Fisher information inequalities, it is demonstrated that SR cannot improve locally optimal processing under the usual conditions. However, outside these restrictive conditions of weak signal and large sample size, improvement by addition of noise through SR can be achieved, and becomes an attractive option for nonlinear signal processing. The analysis in the paper has focused on the simplest case of additive white noise as an essential reference, and an interesting extension for future work is to examine the affect of considering different forms of colored noise [15,31,32].

Methods

Under the assumption of weak signals, the Taylor expansion of the noise PDF is utilized in Eqs. (4), (5), (16) and (21). The Cauchy-Schwarz inequality is extensively used in Eqs. (6), (13), (17), (19) and (26).

Author Contributions

Conceived and designed the experiments: FD. Performed the experiments: FD FCB DA. Analyzed the data: FD FCB DA. Contributed reagents/materials/analysis tools: FCB DA. Wrote the paper: FD. Proofreading: FCB DA.

References

1. Cover TM, Thomas JA (1991) Elements of Information Theory. New York: Wiley.
2. Kay S (1998) Fundamentals of Statistical Signal Processing. N.J.: Prentice-Hall, Englewood Cliffs.
3. Capon J (1961) On the asymptotic efficiency of locally optimum detectors. IRE Transactions on Information Theory IT-7: 67–71.
4. Poor HV, Thomas JB (1978) Locally optimum detection of discrete-time stochastic signals in nongaussian noise. Journal of the Acoustical Society of America 63: 75–80.
5. Lu NH, Eisenstein B (1981) Detection of weak signals in non-gaussian noise. IEEE Transactions on Information Theory IT-27: 755–771.
6. Kassam SA (1988) Signal Detection in Non-Gaussian Noise. New York: Springer-Verlag.
7. Song I, Kassam SA (1990) Locally optimum detection of signals in a generalized observation model: The random signal case. IEEE Transactions on Information Theory 36: 516–530.
8. Benzi R, Sutera A, Vulpiani A (1981) The mechanism of stochastic resonance. Journal of Physics A: Mathematical and General 14: L453–L457.
9. Zozor S, Amblard PO (2003) Stochastic resonance in locally optimal detectors. IEEE Transactions on Signal Processing 51: 3177–3181.
10. Chapeau-Blondeau F, Godivier X (1997) Theory of stochastic resonance in signal transimission by static nonlinear systems. Physical Review E 55: 1478–1495.
11. Chapeau-Blondeau F, Rousseau D (2009) Raising the noise to improve performance in optimal processing. Journal of Statistical Mechanics: Theory and Experiment 1: no. P01003.
12. Dykman MI, McClintock PVE (1998) What can stochastic resonance do? Nature 391: 344.
13. Collins JJ, Chow CC, Imhoff TT (1995) Stochastic resonance without tuning. Nature 376: 236–238.
14. Bezrukov SM, Vodyanoy I (1997) Stochastic resonance in non-dynamical systems without response thresholds. Nature 385: 319–321.
15. Gammaitoni L, Hänggi P, Jung P, Marchesoni F (1998) Stochastic resonance. Reviews of Modern Physics 70: 233–287.
16. Wiesenfeld K, Moss F (1995) Stochastic resonance and the benefits of noise: from ice ages to crayfish and SQUIDs. Nature 373: 33–36.
17. Zaikin AA, Kürths J, Schimansky-Geier L (2000) Doubly stochastic resonance. Physical Review E 85: 227–231.
18. Perc M, Gosaka M, Kralj S (2008) Stochastic resonance in soft matter systems: combined effects of static and dynamic disorder. Soft Matter 4: 1861–1870.

19. Gosaka M, Perc M, Kralj S (2012) The impact of static disorder on vibrational resonance in a ferroelectric liquid crystal. Molecular Crystals & Liquid Crystals 553: 13–20.
20. McDonnell MD, Abbott D (2009) What is stochastic resonance? definitions, misconceptions, debates, and its relevance to biology. PLoS Computational Biology 5: no. e1000348.
21. Perc M (2007) Stochastic resonance on excitable small-world networks via a pacemaker. Physical Review E 76: no. 066203.
22. Perc M (2008) Stochastic resonance on weakly paced scale-free networks. Physical Review E 78: no. 036105.
23. Wang Q, Perc M, Duan Z, Chen G (2009) Delay-induced multiple stochastic resonances on scalefree neuronal networks. Chaos 19: no. 023112.
24. Gan C, Perc M, Wang Q (2010) Delay-aided stochastic multiresonances on scale-free fitzhughnagumo neuronal networks. Chinese Physical B 19: no. 040508.
25. Stocks NG (2000) Suprathreshold stochastic resonance in multilevel threshold systems. Physical Review Letters 84: 2310–2313.
26. Kay S (2000) Can detectabilty be improved by adding noise? IEEE Signal Processing Letters 7: 8–10.
27. Ward LM, MacLean SE, Kirschner A (2010) Stochastic resonance modulates neural synchronization within and between cortical sources. PLoS ONE 5: no. e14371.
28. Duan F, Chapeau-Blondeau F, Abbott D (2011) Fisher-information condition for enhanced signal detection via stochastic resonance. Physical Review E 84: no. 051107.
29. Chen H, Varshney PK, Michels JH, Kay SM (2007) Theory of the stochastic resonance effect in signal detection: Part I–fixed detectors. IEEE Transactions on Signal Processing 55: 3172–3184.
30. Greenwood PE, Müller UU, Ward LM (2004) Soft threshold stochastic resonance. Physical Review E 70: no. 051110.
31. Zeng L, Li J (2011) Residual aperiodic stochastic resonance in Lévy noise. Fluctuation and Noise Letters 10: 223–229.
32. Hänggi P, Jung P, Zerbe C, Moss F (1993) Can colored noise improve stochastic resonance? Journal of Statistical Physics 70: 25–47.
33. DeWeese M, Bialek W (1995) Information flow in sensory neurons. Nuovo Cimento 17D: 733–741.
34. Brown LD (1982) A proof of the central limit theorem motivated by the cramér-rao inequality. Statistics and Probability: Essays in Honor of C R Rao G. Kallianpur, PR. Krishnaiah, JK. Ghosh, eds. North-Holland Publishing Company. pp 141–148.