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

Fabing Duan1*, François Chapeau-Blondeau2, Derek Abbott3

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.

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 Cramer-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

\[ y(t) = g[x(t)], \]  

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 ≤ |s(t)| ≤ 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 \( \sigma_z \). It is assumed that \( g \) has zero mean under \( f_z \), i.e., \( \int_{-\infty}^{\infty} g(x)f_z(x)dx = \mathbb{E}[g(x)] = 0 \), 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 \( \Delta B \) around \( 1/T \), this is

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

with \( \Delta t \) 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

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

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

In the case of \( A \rightarrow 0 \), we have a Taylor expansion of the expectation at a fixed time \( t \) as

\[
E[y(t)] = \int_{-\infty}^{\infty} g(x)f_z(x-s)dx \approx \int_{-\infty}^{\infty} g(x)f_z(x-s)f_z(x)dx \tag{4}
\]

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

\[
\var{y(t)} = \mathbb{E}[y^2(t)] - \mathbb{E}[y(t)]^2 \approx \mathbb{E}[y^2(t)] - s^2(t)\mathbb{E}[g'(x)]
\]

\[
\approx \int_{-\infty}^{\infty} g'(x) f_z(x-s) f_z(x) dx 
\tag{5}
\]

where \( s(t) \) \( \int_{-\infty}^{\infty} g'(x) f_z(x-s) f_z(x) dx = 2s(t)\mathbb{E}[g(x)g'(x)] + s^2(t)\mathbb{E}[g'(x)], \) compared with \( \mathbb{E}[g^2(x)], \) can be neglected as \( A \rightarrow 0 \) \( 0 < |s(t)| \leq A \) \[2,6\]. 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

\[
R_{out} \approx \left\{ \frac{\langle s(t) \exp[-i2\pi t/T] \rangle^2}{\Delta B \Delta t} \right\} \mathbb{E}[g'(x)] \tag{6}
\]

where the expectation \( \mathbb{E}[(f_z'(x))/(f_z(x))] \) is simply the Fisher information \( I(f_z) \) of the noise PDF \( f_z \) \[2,6\], and the equality occurs as

\[
g(x) = C f_z'(x) f_z(x) \Delta g_{\text{opt}}(x), \tag{7}
\]

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

\[
G = \frac{R_{out}}{R_{in}} \approx \frac{\mathbb{E}[g'(x)]^2}{\mathbb{E}[g^2(x)]} \leq \frac{\mathbb{E}[f_z'(x)]^2}{\mathbb{E}[f_z(x)]^2} = \sigma_z^2 I(f_z) = I(f_{z0}). \tag{8}
\]

with equality achieved when \( g \) takes the locally optimal nonlinearity \( g_{\text{opt}} \) of Eq. (7). Here, for a standardized PDF \( f_{z0} \) with zero mean and unity variance \( \sigma_z^2 = 1 \), the scaled noise \( z(t) = \sigma_z z(t) \) has its PDF \( f_z(z) = f_{z0}(z/\sigma_z)/\sigma_z \) and the Fisher information satisfies \( I(f_z) = I(f_{z0})/\sigma_z^2 \) \[1,3,4\]. It is known that a standardized Gaussian PDF \( f_{z0}(z) = \exp(-z^2/2)/\sqrt{2\pi} \) has the minimal Fisher information \( I(f_{z0}) = 1 \) and any standardized non-Gaussian PDF \( f_{z0} \) has the Fisher information \( I(f_{z0}) > 1 \) \[2\]. It can be seen that, the linear system \( g_1(x) = x \) has its output signal-to-noise ratio \( R_{out} = R_{in} \) 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 \( g_1 \).

**Optimum asymptotic efficacy for signal detection**

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

\[
X_n = \theta_{sn} + z_n, \quad n = 1, 2, \ldots, N, \tag{9}
\]

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 \( \theta_{sn} \) are with the signal strength \( \theta \) \[6\]. For the known signal sequence \( \{s_n, n = 1, 2, \ldots, N\} \), 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., \( P_s^\infty = \lim_{N \to \infty} \sum_{n=1}^{N} s_n^2/N < \infty \) \[6\]. 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

\[
H_0 : f_X(X) = \prod_{n=1}^{N} f_z(X_n), \quad \theta = 0; \tag{10}
\]

\[
H_1 : f_X(X) = \prod_{n=1}^{N} f_z(X_n - \theta_{sn}), \quad \theta > 0.
\]

Consider a generalized correlation detector

\[
T_G(X) = \sum_{n=1}^{N} g(X_n) s_n \sim \mathcal{G}_0 \tag{11}
\]

where the memoryless nonlinearity \( g \) has zero mean under \( f_z \), i.e., \( \mathbb{E}[g(x)] = 0 \) \[6\]. In the asymptotic case of \( \theta \to 0 \) and \( N \to \infty \), the test statistic \( T_G \) according to the central limit theorem, converges to a Gaussian distribution with mean \( \mathbb{E}[T_G[H_0]] = 0 \) and variance \( \text{var}[T_G[H_0]] \approx N P_s^\infty \mathbb{E}[g^2(x)] \) under the null hypothesis \( H_0 \) \[6\]. Using Eqs. (4) and (5), \( T_G \) is asymptotically Gaussian with
Consider a generalized energy detector over the linear correlation detector when locally optimal detector established by the Taylor expansion of the benchmark [5,6], the asymptotic relative efficiency with equality being achieved when $g = g_{\text{opt}}$ 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 $\ln \left[ \frac{\pi_{n-1} f(x_n - \theta_0)}{\pi_{n-1} f(x_n)} \right] \approx \sum_{n=1}^{N} g_{\text{opt}}(X_n) \theta_0 (C = -1)$ in terms of the generalized Neyman-Pearson lemma [2,6].

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

$$\frac{\xi_{\text{GC}}}{\xi_{\text{LC}}} = \frac{E[g^2(x)]}{E[g^2(x)]} \leq \sigma_x^2 I(f) = I(f_0),$$

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]

$$H_0 : f(x) = \prod_{n=1}^{N} f_s(x_n), \text{ for } \theta = 0;$$

$$H_1 : f(x) = \int_{-\infty}^{\infty} \prod_{n=1}^{N} f(x_n - \theta_0) f_s(s_n) ds_n, \text{ for } \theta > 0,$$

for determining whether the random signal is present or not.

Consider a generalized energy detector

$$T_{\text{GE}}(X) = \sum_{n=1}^{N} g(X_n) \frac{H_1}{H_0},$$

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

$$E[T_{\text{GE}}|H_1] = N \int_{-\infty}^{\infty} g(x) \int_{-\infty}^{\infty} f(x - \theta_0) f_s(s) ds dx$$

$$\approx N \int_{-\infty}^{\infty} g(x) \int_{-\infty}^{\infty} \left[ f(x) - \theta S f_s(x) + \frac{\theta^2}{2} f''(s) \right] f_s(s) ds ds dx$$

$$\approx N \int_{-\infty}^{\infty} g(x) f''(s) ds ds = N \theta^2 \frac{E[g^2(x)]}{2}.$$

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

$$\zeta_{\text{GE}} = \lim_{N \to \infty} \left( \frac{E[T_{\text{GE}}|H_1]}{\text{var}[T_{\text{GE}}|H_1]^{1/2}} \right)^2 = \frac{1}{4} \frac{E[g^2(x)]}{E[g^2(x)]} \leq \frac{1}{4} \frac{E[f^2(x) \theta^2]}{E[f^2(x)]} = 1 = I(f).$$

where $\theta^2$ is treated as the signal strength parameter and $I(f)$ is the second order Fisher information [6,7]. It is noted that the equality of Eq. (17) is achieved as $g(x) = g_{\text{opt}}(x) = Cf''(s)/f_s$ 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 $\zeta_{\text{GE}}$ [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 $\sigma^2_t$, the zero mean and the upper bound $\Lambda (0 \leq |s(t)| \leq \Lambda)$. 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]

$$\rho_{s,x} = \frac{E[s(t) x(t)]}{\sigma_s \sqrt{E[x^2(t)]}} = \frac{\sigma_z}{\sigma_s} \approx \frac{\sigma_z}{\sigma_s} \frac{E[g^2(x)]}{\sqrt{E[g^2(x)]}} \leq \frac{\sqrt{I(f)}}{I(f_0)},$$

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

$$\rho_{s,y} \approx \frac{\sigma_z}{\sigma_s} \frac{E[g^2(x)]}{\sqrt{E[g^2(x)]}} \leq \frac{\sqrt{I(f)}}{I(f_0)},$$

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

$$G_p \leq \frac{\rho_{s,y}}{\sigma_z} \approx \frac{\sigma_z}{\sigma_s} \frac{E[g^2(x)]}{\sqrt{E[g^2(x)]}} \leq \frac{\sqrt{I(f)}}{I(f_0)}.$$

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 $\theta$. As the upper bound $A \to 0 (0 \leq |s_n| \leq A)$, the Cramer-Rao inequality indicates that the mean squared error of any unbiased estimator of the parameter $\theta$ is lower bounded by the reciprocal of the Fisher information [1,2] given by

$$I(\theta) = \sum_{n=1}^{N} \left[ \frac{\partial}{\partial \theta} \ln f_s(x_n - s_n) \right]^2$$

$$\approx \sum_{n=1}^{N} \left[ \frac{df_s(z_n)}{f_s(z_n)} \right]^2 f_s(z_n) \left[ d(-z_n - s_n) \right]^2 \left[ \frac{- \partial s_n}{\partial \theta} \right]^2$$

$$= I(f_0) \sum_{n=1}^{N} \left[ \frac{\partial s_n}{\partial \theta} \right]^2,$$

which indicates that the minimum mean square error of any unbiased estimator is also determined by the Fisher information $I(f_0)$ of a distribution, as $\sum_{n=1}^{N} \left[ \frac{\partial s_n}{\partial \theta} \right]^2$ 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 $I(f_i)$ 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

$$f_g(x) = \frac{c_1}{\sigma_z^2} \exp\left(-\frac{x}{\sigma_z}\right)^2,$$

where $c_1 = \frac{1}{\sqrt{2\pi}} \frac{1}{\Gamma(3\alpha^{-1})} \frac{1}{\Gamma(\alpha^{-1})}$ for a rate of exponential decay parameter $\alpha > 0$ [2,6]. The corresponding locally optimal nonlinearity is $g_{\text{opt}}(x) = |x|^{\alpha^{-1}} \text{sign}(x)$ and the output-input signal-to-noise ratio gain in Eq. (8) is $G = I(f_g) = \alpha^2 \Gamma(3\alpha^{-1}) \Gamma(2-\alpha^{-1}) \Gamma(\alpha^{-1}) \left(\frac{1}{\alpha}\right)^{\frac{2}{3}}$. For comparison, we also operate the sign nonlinearity $g_\text{sign}(x) = \text{sign}(x)$ and the linear system $g_l(x) = x$ in the generalized Gaussian noise. The output-input signal-to-noise ratio gain in Eq. (8) of $g_\text{sign}$ is $G = 4\alpha^2 \Gamma(3\alpha^{-1}) \left(\frac{1}{\alpha}\right)^{\frac{2}{3}}$ (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 $g_\text{sign}$ attains that of the locally optimal nonlinearity of $g_{\text{opt}}$. This is because, the nonlinearity $g_\text{sign}$ is just the locally optimal nonlinearity for Laplacian noise ($\alpha = 1$), and the Fisher information limit $I(f_{g_{\text{opt}}}) = 2$ is achieved. Likewise, for Gaussian noise ($\alpha = 2$), the linear system $g_l$ is optimal and the output-input SNR gain $G = I(f_{g_{\text{opt}}}) = I$. 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).

**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 [3–32]. 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 $\nu(t)$, independent of the initial noise $z(t)$ and the signal $s(t)$, to $x(t)$. Then, the updated data $\tilde{x}(t) = s(t) + z(t) + \nu(t) = s(t) + \nu(t)$. Here, the composite noise $\nu(t)$ has a convolved PDF

$$f_\nu(x) = \int_{-\infty}^{\infty} f_z(x-u)f_\nu(u)du,$$

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

$$I(f_z) \leq \min(I(f_c),I(f_i)), \quad (24)$$

$$I_s(f_z) \leq \min(I_s(f_c),I_s(f_i)). \quad (25)$$

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

$$I(f_z) = \int f'_z(x)^2 dx \int f_z(x) dx$$

$$\leq \int \left( \int f'_z(x-u)f_\nu(u)du \right)^2 f_z(x) dx \int f_\nu(u)^2 du$$

$$= \int \left( f'_z(x) \right)^2 f_z(x) dx \int f_\nu(u)^2 du$$

Similarly, substituting $f'_z(x) = \int_{-\infty}^{\infty} f'_z(x-u)f_\nu(u)du$ into Eq. (26), we also obtain $I_s(f_z) \leq I_s(f_i)$ 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 outside restrictive conditions, where a suboptimal detector is adopted for Gaussian noise. Consider a generalized correlation

![Figure 1. The output-input signal-to-noise ratio gain](https://www.plosone.org)
detector of Eq. (11) based on the dead-zone limiter nonlinearity

\[ g_{DZ}(x) = \begin{cases} 
-1 & \text{for } x < -c, \\
0 & \text{for } -c \leq x \leq c, \\
+1 & \text{for } x > c,
\end{cases} \quad (27) \]

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

\[ \xi_{GC} = \frac{1}{c^2} \left( \frac{c}{\sigma_z} \right)^2 \frac{2F_{G}(c/\sigma_z)}{1 - F_{G}(c/\sigma_z)}, \quad (28) \]

where \( F_{G} \) is the cumulative distribution function of the standardized generalized Gaussian noise PDF \( f_{G} \) [28]. For a fixed response threshold \( c (c = 1 \text{ without loss of generality}) \), we plot the the normalized asymptotic efficacy \( \xi_{GC} \) (solid line) of the dead-zone limiter nonlinearity \( g_{DZ} \) as a function of the RMS amplitude \( \sigma_z \) of Gaussian noise \( (\sigma_z = 2) \), as shown in Fig. 2. It is clearly seen in Fig. 2 that the SR effect appears, and \( \xi_{GC} \) achieves its maximum \( \xi_{GC} = 1.1512 \) at a non-zero level of \( \sigma_z^2 = 0.6098 \). When the original Gaussian noise RMS \( \sigma_z < \sigma_z^2 = 0.6098 \), we can add independent Gaussian noise \( \nu(t) \) with its RMS amplitude \( \sigma_z = \sqrt{\sigma_z^2 - \sigma_z^2} \) to increase \( \xi_{GC} \) 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**


