Stochastic Resonance with Colored Noise for Neural Signal Detection

Fabing Duan1*, François Chapeau-Blondeau2, Derek Abbott3

1 Institute of Complexity Science, Qingdao University, Qingdao, P. R. China, 2 Laboratoire d’Ingénierie des Systèmes Automatisés, Université d’Angers, Angers, France, 3 Centre for Biomedical Engineering and School of Electrical & Electronic Engineering, The University of Adelaide, Adelaide, Southern Australia, Australia

Abstract
We analyze signal detection with nonlinear test statistics in the presence of colored noise. In the limits of small signal and weak noise correlation, the optimal test statistic and its performance are derived under general conditions, especially concerning the type of noise. We also analyze, for a threshold nonlinearity—a key component of a neural model, the conditions for noise-enhanced performance, establishing that colored noise is superior to white noise for detection. For a parallel array of nonlinear elements, approximating neurons, we demonstrate even broader conditions allowing noise-enhanced detection, via a form of suprathreshold stochastic resonance.

Introduction
Stochastic resonance has emerged as a significant statistical phenomenon where the presence of noise is beneficial for signal and information processing in both man-made and natural systems [1–11]. The excitable FitzHugh–Nagumo (FHN) neuron model has been discussed for exploring the functional role of noise in neural coding of sensory information [12]. Following this, the milestone concept of aperiodic stochastic resonance using the FHN neuron model [13] stimulated a number of interesting investigations in sensory biology [3,7,14,15] and physiological experiments [6,8,16–20]. Due to the character of activity in the nervous system, the neuron coding strategy based on stochastic resonance is also found in threshold (level-crossing) [21–23] and nervous system, the neuron coding strategy based on stochastic resonance has manifested its potentiality.

In many practical situations, the idealization of white noise is never exactly realized [2,5]. Consequently, the effect of colored noise on stochastic resonance has been investigated using the measure of output signal-to-noise ratio of a periodic signal [2,5,16,41–43]. Although the suppression of stochastic resonance with increasing noise correlation time was demonstrated [2,5,41–43], it is interesting to note that, under certain circumstances, colored noise can be superior to white noise for enhancing the response of a nonlinear system to a weak signal [16,44]. In the field of signal detection, the employment of noise to enhance signal detectability also becomes a possible option [45–55]. However, in most of these studies, the observed noise samples are often assumed to be independent. Colored noise for signal detection [56–60] is not adequately investigated in the context of stochastic resonance. In this article, we focus on the weak signal detection problem with the beneficial role of additive colored noise in threshold neurons. Because of the “all-and-none” character of nerve activity [61], the problem of threshold-based neural signal detection can be considered as a statistical binary hypothesis test [7,23,27]. In this situation, explicit expressions for the maximum asymptotic detection efficacy are derived for a given transfer function of neuron model. We prove that colored noise that arises from a moving-average model is superior to white noise in improving the detection efficacy of neurons. It is illustratively shown that, for a single neuron with a signum threshold nonlinearity, the possibility of noise-enhanced detection only holds for non-scaled noise. For scaled noise, the effect of noise-enhanced detection does not occur in a single neuron model. However, when we tune the internal noise components of a parallel array of threshold neurons, it is observed that the constructive role of noise comes into play again in improving the signal detection efficacy, wherein suprathreshold stochastic resonance manifests its potentiality.
Results

Detection model

Consider the detection problem formulated as a binary hypothesis test [23,60,62]

$$H_0 : \mathbf{x} = \mathbf{z},$$
$$H_1 : \mathbf{x} = \mathbf{z} + \theta \mathbf{s}.$$  \hspace{1cm} (1)

Under hypothesis $H_0$, the observation vector $\mathbf{x} = [x_1, x_2, \cdots, x_N]^T$ consists of noise $\mathbf{z} = [z_1, z_2, \cdots, z_N]^T$ only, and under hypothesis $H_1$ it consists of noise $\mathbf{z}$ and known signal $\mathbf{s} = [s_1, s_2, \cdots, s_N]^T$ with its strength $\theta$. There exists a finite bound $U_j$ such that $0 \leq |s_n| \leq U_j$, and the asymptotic average signal power satisfies $0 < \rho_j = \lim_{{N \to \infty}} \mathbf{s}^T \mathbf{s}/N < \infty$ [56–58,60,62]. Next, the test statistic $T(\mathbf{x})$ is compared with a decision threshold $\gamma$ to decide the hypotheses, as

$$T(\mathbf{x}) = \mathbf{c}^T g(\mathbf{x}) \geq \gamma_{H_0},$$  \hspace{1cm} (2)

where the coefficient vector $\mathbf{c} = [c_1, c_2, \cdots, c_N]^T$ is associated with the function $g(\mathbf{x})$ to form $T(\mathbf{x})$.

Assume the $N$-dimensional probability distribution $f(\mathbf{z})$ of noise $\mathbf{z}$ and zero-mean vector of $E[\mathbf{g}(\mathbf{z})] = f(\mathbf{z})(\mathbf{z}) d\mathbf{z}$ (for a shift in mean) [58,60]. Then, for a large sample size $N$ of observation vector $\mathbf{x}$, the test statistic $T(\mathbf{x})$ has zero-mean and asymptotic variance

$$\text{var}(T(\mathbf{x}))[H_0] = \mathbf{c}^T E[g(\mathbf{z})g(\mathbf{z})^T] \mathbf{c},$$

under hypothesis $H_0$. Furthermore, for weak signals ($\theta \to 0$) and under hypothesis $H_1$, $g(\mathbf{x})$ can be expanded to the first-order

$$g(\mathbf{x}) = g(\mathbf{z} + \theta \mathbf{s}) \approx g(\mathbf{z}) + \theta \frac{\partial g(\mathbf{z})}{\partial \mathbf{z}} \mathbf{s}.$$  \hspace{1cm} (4)

Then, the characteristics of $T(\mathbf{x})$ under $H_1$, up to the first-order in $\theta$, can be calculated as

$$\text{E}(T(\mathbf{x})[H_1]) \approx \theta \mathbf{c}^T E \left[ \frac{\partial g(\mathbf{z})}{\partial \mathbf{z}} \mathbf{s} \right], \text{ var}(T(\mathbf{x})[H_1]) \approx \text{var}(T(\mathbf{x})[H_0]).$$  \hspace{1cm} (5)

Under both hypotheses $H_0$ and $H_1$, the test statistic $T(\mathbf{x})$, according to the central limit theorem, converges to a Gaussian distribution. Thus, the binary hypothesis test of Eq. (1) becomes a Gaussian mean-shift detection problem [60,62]. Given the false probability, the detection probability is a monotonically increasing function of the detection efficacy $\xi(T)$ [60,62] given by

$$\xi(T) = \lim_{{N \to \infty}} \left( \frac{\frac{\text{E}(T(\mathbf{x})[H_1])}{\text{E}(T(\mathbf{x})[H_0])}^2}{\text{var}(T(\mathbf{x})[H_0])} \right)^2$$
$$= \left( \frac{\mathbf{c}^T E[g(\mathbf{z})g(\mathbf{z})^T] \mathbf{c}}{\mathbf{c}^T E[g(\mathbf{z})^2] \mathbf{c}} \right)^2$$
$$= \left( \frac{\mathbf{c}^T E[g(\mathbf{z})^2] \mathbf{c}}{\mathbf{c}^T E[g(\mathbf{z})^2] \mathbf{c}} \right)^2$$
$$= \left( \frac{\mathbf{c}^T E[g(\mathbf{z})^2] \mathbf{c}}{\mathbf{c}^T E[g(\mathbf{z})^2] \mathbf{c}} \right)^2$$
$$\leq \mathbf{s}^T \mathbf{J} \mathbf{s},$$

where the Cauchy-Schwarz inequality yields

$$\left( \mathbf{c}^T E[g(\mathbf{z})^2] \left( \frac{\partial \ln f(\mathbf{z})}{\partial \mathbf{z}} \right)^T \mathbf{s} \right)^2 \leq \mathbf{c}^T E[g(\mathbf{z})^2] \left( \frac{\partial \ln f(\mathbf{z})}{\partial \mathbf{z}} \right)^2 \mathbf{c} \mathbf{s}^T \mathbf{J} \mathbf{s},$$  \hspace{1cm} (7)

with the Fisher information matrix $\mathbf{J} = E[\left( \frac{\partial \ln f(\mathbf{z})}{\partial \mathbf{z}} \right)^2]$. Note that the equality of Eq. (6) is satisfied by the locally optimum nonlinearity

$$g_{\text{opt}}(\mathbf{z}) = C \frac{\partial \ln f(\mathbf{z})}{\partial \mathbf{z}},$$  \hspace{1cm} (8)

for an arbitrary constant $C$.

However, a complete closed-form noise distribution $f(\mathbf{z})$ may be unavailable, especially in unknown noisy circumstances [56–58,60,62], which makes the nonlinearity of Eq. (8) difficult or too complex to implement. Thus, there may be compelling reasons for considering the given function $g$ with an easily implemented feature. In this case, the detection efficacy in Eq. (6) can be maximized as

$$\xi(T) = \left( \frac{\mathbf{c}^T E[g(\mathbf{z})^2] \mathbf{c}}{\mathbf{c}^T E[g(\mathbf{z})^2] \mathbf{c}} \right)^2$$
$$= \left( \frac{(\mathbf{L}^T \mathbf{c})^T (\mathbf{L}^{-1} E[g(\mathbf{z})^2] \mathbf{c})}{(\mathbf{L}^T \mathbf{c})^T (\mathbf{L}^T \mathbf{c})} \right)^2$$
$$\leq \mathbf{s}^T \mathbf{J} \mathbf{s},$$

with the Cholesky decomposition of the variance matrix $\mathbf{V} = E[g(\mathbf{z})^2] = \mathbf{L} \mathbf{L}^T$ and by optimally choosing the coefficient vector $\mathbf{c}_{\text{opt}} = \mathbf{c} E[g(\mathbf{z})^2]^{-1} E[g(\mathbf{z})^2] \mathbf{s}$ for an arbitrary constant $\kappa$.

Colored noise

Consider a useful colored noise model of the first-order moving-average [56,59] as

$$z_i = \rho_1 w_{i-1} + \omega_i + \rho_2 w_{i+1},$$  \hspace{1cm} (10)

where the correlation coefficients are $\rho_{1,2}$ and $\mathbf{w} = [w_1, w_2, \cdots, w_N]^T$ is an independent identically distributed (i.i.d.) random vector. For small values of $\rho_{1,2}$ ($|\rho_{1,2}| < 1$), the dependence among noise samples $z_i$ will be weak [56,59]. Here, we assume $w_i$ has an univariate distribution $f_w(w)$ that is symmetric about the origin. We also assume the memoryless nonlinearity $g(\mathbf{z}) = [g(z_1), g(z_2), \cdots, g(z_N)]^T$ to be odd symmetric about the origin. Then, up to first order in small correlation coefficients $\rho_{1,2}$, we can expand $g(z_i)$ as

$$g(z_i) \approx g(w_i) + (\rho_1 w_{i-1} + \rho_2 w_{i+1}) g'(w_i),$$  \hspace{1cm} (11)

$$g'(z_i) \approx g'(w_i) + (\rho_1 w_{i-1} + \rho_2 w_{i+1}) g''(w_i),$$  \hspace{1cm} (12)

and obtain expectations

$$E[g(z_i)] \approx E[g'(w_i)].$$  \hspace{1cm} (13)
E[g^2(z_i)] \approx E[g^2(w)], \quad (14)

E[g(z_i)g(z_{i+1})] \approx (\rho_1 + \rho_2)E[wg(w)]E[g'(w)]. \quad (15)

Therefore, we have the expectation matrix

\[
E\left[ \frac{\delta g(z)}{\delta z} \right] \approx E[g'(w)] I, \quad (16)
\]

with the unit matrix I, and the variance matrix \( V = E[g(z)g(z)^T] \) has elements

\[
V_{ij} = E[g^2(w)], \quad (17)
\]

\[
V_{i,i+1} = V_{i+1,i} = (\rho_1 + \rho_2)E[wg(w)]E[g'(w)], \quad (18)
\]

for \( i = 1, 2, \ldots, N \). Then, based on Eq. (9), the normalized detection efficacy \( e(g, \rho) \) can be computed as

\[
e(g, \rho) = \frac{\ddot{z}(T)}{NP_s} \approx \frac{s^T E[\dot{z}(z)/\dot{z}] V^{-1} E[\dot{z}(z)/\dot{z}] s}{s^T s} = E^2[g'(w)] \frac{s^T V^{-1} s}{s^T s} \leq \frac{E^2[g'(w)]}{\lambda_{\text{min}}}, \quad (19)
\]

Here, when the equality of Eq. (19) is achieved, the signal \( s \) is the corresponding eigenvector to the minimum eigenvalue \( \lambda_{\text{min}} \) of the matrix \( V \). It is known that the eigenvalues of the matrix \( V \) are [62]

\[
\lambda_i = V_{i,i} + 2V_{i,i+1} \cos \left( \frac{\pi i}{N+1} \right) = E[g^2(w)] + 2(\rho_1 + \rho_2)E[wg(w)]E[g'(w)] \cos \left( \frac{\pi i}{N+1} \right), \quad (20)
\]

corresponding to eigenvectors \( v_k = \sin(\pi k/(N+1)) \) for \( k = 1, 2, \ldots, N \). Here, as the nonlinearity \( g \) is assumed to be odd, it is then found that \( E[wg(w)] \geq 0 \) and \( E[g'(w)] \geq 0 \). Therefore, if \( \rho_1 + \rho_2 < 0 \) and for a large sample size \( N \), we take \( i = 1 \) and \( \lambda_{\text{min}} = \lambda_1 \approx E[g^2(w)] + 2(\rho_1 + \rho_2)E[wg(w)]E[g'(w)]. \) Otherwise, we choose \( \lambda_{\text{min}} = \lambda_N \approx E[g^2(w)] - 2(\rho_1 + \rho_2)E[wg(w)]E[g'(w)]. \) An illustration of the eigenvector \( v_N \) is shown in Fig. 1 for \( N = 200 \). In this way, by optimally choosing the input signal (eigenvector) \( s = v_1 \) (\( v_N \)), the maximum efficacy \( e(g, \rho) \) can be calculated as

\[
e_{\text{max}}(g, \rho) = \max_s e(g, \rho) = \frac{E^2[g'(w)]}{V_{i,i} - 2V_{i,i+1}} \quad (21)
\]

Since, from its definition in Eq. (6), the efficacy \( e(g, \rho) \) is non-negative, the denominator in Eq. (21) must satisfy

\[
E[g^2(w)] - 2|\rho_1 + \rho_2|E[wg(w)]E[g'(w)] > 0. \quad (22)
\]

In order to validate Eq. (22), we use the Cauchy-Schwarz inequality to yield

\[
E^2[wg(w)] \leq E[w^2]E[g^2(w)] = \sigma_w^2 E[g^2(w)], \quad (23)
\]

\[
E^2[g'(w)] \leq E[\left( \frac{f'_w}{f_w} \right)^2]E[g^2(w)] = J(f_w)E[g^2(w)], \quad (24)
\]

with the Fisher information quantity \( J(f_w) = E[\left( \frac{f'_w}{f_w} \right)^2] \) and the variance \( \sigma_w^2 \) of noise distribution of \( f_w \) [56]. Thus, we find

\[
e[wg(w)] E[g'(w)] \leq \sqrt{\sigma_w^2 J(f_w)} E[g^2(w)]. \quad (25)
\]

Substituting Eq. (25) into Eq. (22) and noting

\[
\sigma_w^2 J(f_w) = E[w^2]E[\left( \frac{f'_w}{f_w} \right)^2] \geq E\left[ \left( \frac{f'_w}{f_w} \right)^2 \right] = 1, \quad (26)
\]

we have

\[
|\rho_1 + \rho_2| \leq \frac{1}{\sqrt{2}} \frac{\lambda_{\text{min}}}{J(f_w)} \leq \frac{1}{2}. \quad (27)
\]

Since we assume \( |\rho_{1,2}| < 1 \), the inequalities of Eqs. (27) and (22) can be satisfied, and the detector efficacy in Eq. (21) will be theoretically analyzed in the following.

For white noise vector \( z \) with zero correlation coefficients \( \rho_{1,2} = 0 \), the detection efficacy \( e(g, \rho) \) in Eq. (21) satisfies

\[
e(g, \rho) = \frac{E^2[g'(w)]}{E[g^2(w)]} \leq e_{\text{max}}(g, \rho). \quad (28)
\]

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Eigenvector \( v_N \). An illustration of the eigenvector \( v_N \) of the variance matrix \( V(N = 200) \).}
\end{figure}

doi:10.1371/journal.pone.0091345.g001
Thus, for a given function $g$, colored noise is superior to white noise in enhancing the detection efficacy, at a cost of optimally matching the input signal with the eigenvector $v_i$ of covariance matrix $V$.

**Stochastic resonance in threshold-based neurons**

We will illustratively show the possibilities of noise-enhanced detection in threshold-based neurons. The classical McCulloch-Pitts threshold neuron has the form

$$g(x) = \begin{cases} 1, x > \ell, \\ 0, x \leq \ell, \end{cases}$$ 

(29)

with the response threshold $\ell$. It is seen that $g$ can be expressed as a function of $x$ in terms of the signum (sign) function as $g(x) = \frac{1}{2} \text{sign}(x-\ell) + \frac{1}{2}$. Since the constant factor $1/2$ does not affect the detection efficacy of the transfer function $g$, then we focus on the signum function

$$g(x) = \text{sign}(x),$$

(30)

with response threshold $\ell = 0$ in the following parts. Here, the signum function $g$ is not continuous at $x = 0$, but has its derivative $g'(x) = 2\delta(x)$ for any $x$ [60].

For the colored noise model of Eq. (10), the correlation coefficient $\rho_2 = 0$ indicates the noise sequence $z_i$ is a causal process that can be physically realized. Here, we assume $p_1 = \mu$ (|$\mu$| < 1) and $\rho_2 = 0$, and show the possibility of stochastic resonance in the physically realizable noise environment. First, consider scaled noise $w(t) = \sigma_{w_0}w(t)$ that has the distribution $f_{w_0}(w) = f_{w_0}(w/\sigma_{w_0})/\sigma_{w_0}$ [23,60]. Here, $w_0(t)$ has a standardized distribution $f_{w_0}$ with unity variance $\sigma_{w_0}^2 = 1$. Thus, based on Eq. (21), the absolute moment is

$$E[wg(w)] = E[w \text{sign}(w)] = E[|w|] = \sigma_{w_0}E[w_0]|w_0|,$$

(31)

where the operator $E[w_0]$ is $\int f_{w_0}(w_0)dw_0$. Thus, for the signum function $g$, the detection efficacy of Eq. (21) can be expressed as

$$e_{\text{max}}(g, \rho) = \frac{4\sigma_{w_0}^2 g(0)}{1 - 4|\rho|f_{w_0}(0)E[w_0]|w_0|} = \frac{4\sigma_{w_0}^2 g(0)}{1 - 4|\rho|f_{w_0}(0)E[w_0]|w_0|}.$$

(32)

It is seen in Eq. (32) that $e_{\text{max}}(g, \rho)$ is a monotonically decreasing function of noise variance $\sigma_{w_0}^2$, and no noise-enhanced detection effect will occur in such a single neuron model for scaled noise.

We further consider non-scaled Gaussian mixture distribution [5,47,48,60]

$$f_{w}(x) = \frac{1}{2\sqrt{\pi \sigma^2}} \left[ \exp \left( -\frac{(x-\mu)^2}{2\sigma^2} \right) + \exp \left( -\frac{(x+\mu)^2}{2\sigma^2} \right) \right],$$

(33)

where the variance $\sigma^2 = \mu^2 + \epsilon^2$ and parameters $\mu, \epsilon \geq 0$. Then, for the signum function $g$ in Eq. (30), the detection efficacy of Eq. (21) can be computed as

$$e_{\text{max}}(g, \rho) = \frac{2}{1 - 2|\rho| \sqrt{\pi} \int_{-\infty}^{\infty} \exp \left( -\frac{t^2}{2} \right) dt},$$

(34)

where the error function $\text{erf}(x) = 2/\sqrt{\pi} \int_{0}^{x} \exp \left( -t^2 \right) dt$. In Fig. 2, for the correlation coefficient |$\rho$| = 0.2 and different values of $\mu = 0.3, 0.5$ and 1, we show the detection efficacy of Eq. (34) as a function of noise variance $\sigma_{w_0}^2$. For a given non-zero value $\mu$ and as $\epsilon \to 0$, the noise distribution model of Eq. (33) indicates the dichotomous noise [5,48,53]. In this situation, as the signal strength $\theta \to 0$ and $||\theta|/\mu|\ll \mu$, the signum function $g$ will not change its output whether the signal appears or not. Therefore, the test statistics $T(x) = \epsilon^2 g(x)$ in Eq. (2) will be the same value under hypotheses $H_0$ and $H_1$, and the detection efficacy $e_{\text{max}}(g, \rho)$ in Eq. (34) starts from zero. This explanation can be also validated by Eq. (34) as $\epsilon \to 0$ and $\mu$ being fixed, as illustrated in Fig. 2. However, it is clearly seen that, upon increasing the noise variance $\sigma_{w_0}^2 = \epsilon^2 + \mu^2$ (actually increasing $\epsilon$), the noise-enhanced detection effect appears. The smaller the parameter $\mu$ is, the more pronounced the resonant peak of $e_{\text{max}}(g, \rho)$ becomes, as shown in Fig. 2.

Next, an interesting problem is that, for scaled noise, can we observe the noise-enhanced detection effect in threshold-based neurons? The answer to this question is affirmative. Here, we will resort to the constructive role of internal noise for improving the performance of an array of threshold neurons. Let $X = [x_{1m}, x_{2m}, \ldots, x_{Nm}]^T$ be the vector of $N$ observation components at the $m$-th element of receiving array of $M$ identical neurons. In this observation model, $X = z_i + y_i$, under the hypothesis $H_0$. Here, in each neuron element, the $M$ noise terms $y_i$ are assumed to be mutually independent with the same PDF $f_y$ and variance $\sigma_{y}^2$. Then, at the observed time $i$, the array outputs are collected as $g(x_i) = \sum_{m=1}^{M} g(x_{im})/M$, and the test statistics can be reconstructed as $T_{\text{MC}}(x) = \epsilon^2 g(x)$ with $g(x_i) = [g(x_1), g(x_2), \ldots, g(x_N)]^T$. For the colored noise model of Eq. (10) with $p_1 = \rho$ and $p_2 = 0$, we have

$$g(x_i) \approx g(v_i) + \mu w_{i-1} g'(v_i),$$

(35)

$$g'(x_i) \approx g(v_i) + \mu w_{i-1} g'(v_i),$$

(36)

where the composite noise $v_i = w_{i-1} + y_i$ has the convolved distribution $f_{w}(v) = \int f_{w}(v-u)f_{y}(u)du$. Then, we have expectations

$$E_{x}[g'(x_i)] = \frac{1}{M} \sum_{m=1}^{M} E_{x}[g'(x_{im})] \approx E_{x}[g'(v)],$$

(37)

and

$$E_{x}[\frac{\partial g(x)}{\partial x}] \approx E_{x}[g'(v)] I,$$

(38)

with the operator $E_{x}[:]=\int f_{x}(v)dv$. The variance matrix $\nabla = E_{x}[g'(x)g(x)^T]$ has elements

$$\nabla_{ij} = E_{x}[g'(x_i)g'(x_j)]$$

$$= E_{x}\left\{ \frac{1}{M} \sum_{m=1}^{M} \sum_{n=1}^{M} E_{y}[g'(x_{im})g'(x_{jn})] \right\}$$

$$\approx \frac{1}{M} E_{x}[g'(v)] + \frac{(M-1)}{M} E_{w}[E_{y}[g'(w+y)]]$$

(39)
The resonant peaks of \( \epsilon_{\text{max}}(g, \rho) \) are marked by the square (\( M \)), the star (\( \star \)) and the down triangle (\( \downarrow \)) for \( \mu = 0.3, 0.5 \) and 1, respectively. Here, the transfer function \( g(x) = \text{sign}(x) \), and the noise distribution is Gaussian mixture model of Eq. (33).

\[
\max_{\ell}(\text{sign}(x)) = \frac{E_{\text{sign}}[g'(v)]}{\nabla_{x,\ell} - 2|\nabla_{x,\ell}|}.
\]

For instance, we assume the initial Gaussian noise components \( \nu_i \) have the distribution of \( f_{\nu}(x) = \frac{1}{\sqrt{2\pi\sigma_{\nu}^2}} \exp(-\frac{x^2}{2\sigma_{\nu}^2}) \) and the given variance \( \sigma_{\nu}^2 \). The internal noise components of each neuron is assumed to be the uniform random variable \( y_j \) with its distribution \( f_{y}(x) = 1/(2b) \) for \(-b \leq x \leq b\) and zero otherwise. The composite random variables \( v_{ij} \) are distributed by

\[
f_{v}(x) = \frac{\text{erf}(\frac{x+b}{\sqrt{2}m}) - \text{erf}(\frac{x-b}{\sqrt{2}m})}{4b}.
\]

For a given Gaussian noise level \( \sigma_{\nu} = 0.3 \), it is shown in Figs. 3 (a) and (b) that the maximum detection efficacy \( \epsilon_{\text{max}}(g, \rho) \) varies as a function of internal uniform noise level \( b \) for different array sizes

For a given transfer function, we maximize the detection efficacy by optimally choosing the signal waveform. We prove that colored noise is superior to white noise in enhancing the detection efficacy of a parallel array threshold-based neurons.
detection efficacy, at a cost of optimally matching the input signal with the eigenvector of the covariance matrix. Furthermore, we illustrate that, for a single threshold neuron, the possibility of noise-enhanced detection cannot occur in scaled noise, but does appear in a non-scaled Gaussian mixture noise model. Furthermore, for scaled noise, we can test a parallel bundle of neurons with the same response threshold, and recover the positive role of internal noise in enhancing the detection efficacy of the neuron array via the mechanism of suprathreshold stochastic resonance. These results demonstrate that the strategy of exploiting stochastic resonance is still interesting in the case of improving the nonlinear system performance by adding more noise to the signal corrupted by colored noise.

Here, we mainly consider the first-order moving-average noise model of Eq. (10) which is, as we show, amenable to analytical treatment. It is possible to extend the present approach to higher-order moving-average noise models. However, the same analytical treatment maybe no longer feasible. It is also interesting to consider yet other models of colored noise to enhance the detectability of the neuron array. This subject is very promising and currently under study.

It is noted that the detection efficacy of Eqs. (6) and (9) are established under the assumption of weak signal strength $\theta = 0$. We only consider the first-order Taylor expansion of nonlinearities in Eq. (4), because it makes an analytical treatment possible and the corresponding results are rigorous. In practice, most noise distributions are symmetric and the nonlinear characteristics are odd symmetric about the origin. In this case, we can expand the nonlinearity to the second-order terms. The expectation of the second-order term of Taylor expansion of Eq. (4) vanishes and does not affect the conclusion of this paper. However, for unsymmetrical noise distributions and nonlinearities, the high-order terms of Taylor expansion of Eq. (4) are not exactly zero. For this case, we need to numerically observe the effect of high-order terms on the detector performance. It is interesting to compare the present theoretical results of first-order expansion with the numerical results in the further studies.

We also note that these equations of Eqs. (4)–(9) are the extension of white noise [54,56–58,60] to the case of colored noise. Then, we consider a model of colored noise allowing for an analytical evaluation of the detection efficacy in Eqs. (6) and (9). The detection efficacy can also be numerically computed to address other models of colored noise, or to explore broader conditions beyond the weak signal limit. As the signal strength $\theta$ increases, the Taylor expansion of Eq. (4) and the upper bound of Eq. (6) gradually cease to apply. However, based on the present results on weak signal in colored noise, and on [25,26,63] on non-weak signal in Gaussian white noise, it can be expected that noise benefit as reported here will persist with colored noise beyond the small-signal limit.

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

**References**