Extracting hidden weak sinusoidal signal with short duration from noisy data: Analytical theory and computational realization*

Ying Zhang(张英)[1], Zhaoyang Zhang(张朝阳)[1,2], Hong Qian(钱弘)[3], and Gang Hu(胡岗)[1,†]

[1] Department of Physics, Beijing Normal University, Beijing 100875, China
[2] Faculty of Science, Ningbo University, Ningbo 315211, China
[3] Department of Applied Mathematics, University Washington, Seattle, WA 98195, USA

(Received 4 August 2017; revised manuscript received 25 August 2017; published online 10 September 2017)

Signal detection is both a fundamental topic of data science and a great challenge for practical engineering. One of the canonical tasks widely investigated is detecting a sinusoidal signal of known frequency ω with time duration T: $I(t) = A\cos\omega t + \Gamma(t)$, embedded within a stationary noisy data. The most direct, and also believed to be the most efficient, method is to compute the Fourier spectral power at ω: $B = \left| \frac{2}{T}\int_0^T I(t)\,\mathrm{e}^{\mathrm{i}\omega t}\,\mathrm{d}t \right|$. Whether one can out-perform the linear Fourier approach by any other nonlinear processing has attracted great interests but so far without a consensus. Neither a rigorous analytic theory has been offered. We revisit the problem of weak signal, strong noise, and finite data length $T = O(1)$, and propose a signal detection method based on resonant filtering. While we show that the linear approach of resonant filters yield a same signal detection efficiency in the limit of $T \to \infty$, for finite time length $T = O(1)$, our method can improve the signal detection due to the highly nonlinear interactions between various characteristics of a resonant filter in finite time with respect to transient evolution. At the optimal match between the input $I(t)$, the control parameters, and the initial preparation of the filter state, its performance exceeds the above threshold B considerably. Our results are based on a rigorous analysis of Gaussian processes and the conclusions are supported by numerical computations.

Keywords: signal detection, strong noise, stochastic theory, time series analysis

PACS: 05.10.Gg, 05.45.Tp, 02.50.–r **DOI:** 10.1088/1674-1056/26/10/100508

1. Introduction

Signal detections are crucial tasks in various fields.[1–3] In many practical situations, signals may be very weak and short-lived, and they are often contaminated by strong noise. How to extract weak signal from data with a short duration in the presence of strong noisy background turns out to be a significant issue both in theories and in practices.[4–8] We shall revisit the simplest as well as most basic signal-detection problem: Assume a section of sinusoidal signal of duration T, $s(t) = \Pi(t;\tau_1,\tau_2)A\cos\omega_0 t$, $\tau_2 = \tau_1 + T$ (the green curve in the small frame of Fig. 1(a)) is embedded in a strong white noise background $\Gamma(t)$ as shown in Fig. 1(a). Here $\Pi(t;\tau_1,\tau_2)$ is 1 for $\tau_1 \leqslant t \leqslant \tau_2$ and zero otherwise. Throughout the paper, ω_0 and T are assumed to be known, for in many practical cases signals can be emitted by the receivers themselves. Also, we take $\omega_0 = \omega = 0.02$ and $D = 0.5$. Due to extremely small signal-to-noise ratio proportional to A^2/D ($A^2 \ll D$), one can hardly find any trace of signal from the whole input $I(t) = s(t) + \Gamma(t)$ shown as blue curve in Fig. 1(b). A direct way of sinusoidal signal detection is to make Fourier analysis and test different τ for finding the time period with signal $s(t)$. We thus have

$$B(\tau) = B(\omega_0, T, \tau) = \left| \frac{2}{T}\int_\tau^{\tau+T} I(t)\,\mathrm{e}^{i\omega_0 t}\,\mathrm{d}t \right|, \qquad (1)$$

which is expected to give the best distinction of the signal part from the noise background if the τ used in Eq. (1) just fits the detected signal segment of $\tau = \tau_1$. In Fig. 1(c) we plot $B(\tau)$ versus τ for the input $I(t)$ in Fig. 1(b), and it seems that from $B(\tau)$ there is not any detectable feature responding the signal $s(t)$ in Fig. 1(a).

For many years, scientists tried to answer some fundamental as well as practical questions related to the signal detection of Fig. 1, such as, (i) Whether we can enhance the signal-to-noise ratio (SNR) exceeding the input B by any designed means; (ii) How much we can do this if we can. Many methods have been suggested in previous investigations,[9–12] among which the method of stochastic resonance (SR) has been expected to be most hopeful.[9,10] However, despite of the efforts of more than two decades, no satisfactory results have been obtained. First, due to the nonlinearity of SR systems, no exact analysis is possible and, in particular, no convincing answer based on rigorous mathematical–physical computations to the first question has been made. Second though some numerical analysis show enhancement of SNR of SR output out-

*Project supported by the Key Program of the National Natural Science Foundation of China (Grant No. 11135001).
†Corresponding author. E-mail: ganghu@bnu.edu.cn
© 2017 Chinese Physical Society and IOP Publishing Ltd

performing B of input, the enhancement is possible only for the inputs with rather large SNR for which signal $s(t)$ can be easily detected even without SR treatment.[13–15] The practically meaningful problem, whether and how we can enhance the effect of signal detection over the threshold of the input B with weak signals contaminated by strong noises is remaining not answered. Third, all previous investigations have focused on infinitely long time length. The problem of practical significance how to enhance the detection efficiency of the shortly-lived signal from strong noise background has seldom been considered before.

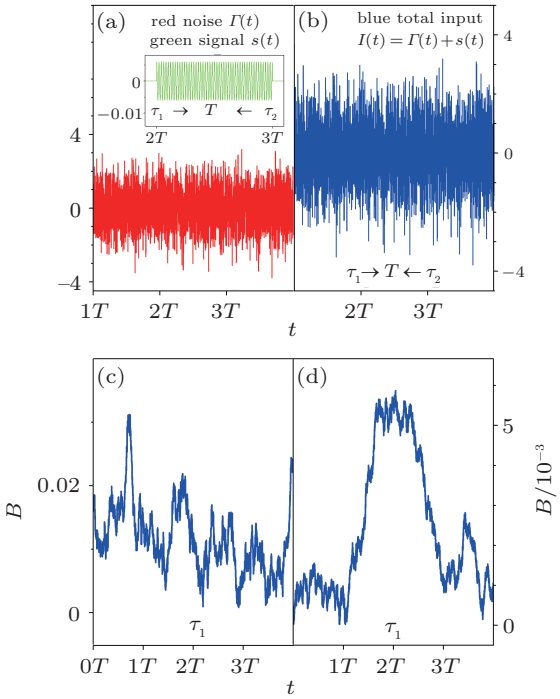


Fig. 1. (color online) Detecting weak and shortly-lived signal from strong noise background. (a) Pure noise time sequence $\Gamma(t)$, D=0.5 (red curve), and signal $s(t) = A\cos(\omega t)$ in the time interval ($\tau_1 = 2T$ to $\tau_2 = 3T$) with $T = 100\pi/\omega$, $\omega = 0.02$, $A = 0.006$ (green curve). (b) Time sequence of weak signal contaminated with strong noise $I(t) = s(t) + \Gamma(t)$ with $s(t)$ and $\Gamma(t)$ given in panel (a), from which one can hardly find any trace of $s(t)$. (c) $B(\tau)$ computed by Eq. (1) for $A = 0.006$ with finite time length $T = 100\pi/\omega$. Note, signal exists in time interval $t = [\tau_1, \tau_2]$, which cannot be observed in the plots. (d) The same as panel (c) with longer signal interval, $T' = \tau_2 - \tau_1 = 10^4\pi/\omega$. $B(\tau)$ during the signal interval is considerably above that in the rest times. Now signal in $[\tau_1, \tau_2]$ interval with $A \approx 0.006$ can be evidently extracted.

The central tasks of the present paper are the following. i) Considering the problem of extracting weak signal from strong noise background; ii) Detecting signal with short life time; iii) Most importantly, applying rigorous computation to analytically answer whether we can and how much we can enhance SNR of outputs from the result of inputs.

For many years scientists have been familiar with the method to use linear resonant filters to amplify the intensity of input signals with certain frequency. It is commonly believed that linear filters can never enhance SNR of outputs over the SNR of the inputs, for all linear filters can equally amplify the signal and the background noise round its resonant frequency due to the linearity of the filters, and then these devices can never enhance the ratio of $\mathrm{SNR_{out}/SNR_{in}}$ over 1. This understanding is valid and exactly proven for any sinusoidal signal with infinite length. It has never been aware so far that this seemly solid conclusion may not been valid in the transient processes of linear filters which treat shortly-lived signal detection. It is found in this work that for finite time inputs of noisy signal interact with system parameters and initial prepared states in a highly nonlinear manner and this nonlinearity provides a possibility to enhance the output SNR exceeding the input SNR.

This paper is organized as follows. In Section 2, we describe our linear filter system and define SNR's for both input and output with finite signal-life-time, and analytically compute SNR of the input. In Section 3, the output SNR is analytically derived too. $\mathrm{SNR_{out}/SNR_{in}} > 1$ in certain parameter regimes is rigorously proven by theoretical computations. And numerical simulations confirm all theoretical results and confirm the conclusion that for an input with weak and shortly-lived signal and strong noise, linear filters can indeed considerably enhance our ability to extract weak signal from strong noise under suitable match among input, filter control parameters and preparation of initial states of filters.

2. Tasks extracting weak signals with short time durations from strong white noises

Considering the simplest as well as the most fundamental problem of sinusoidal signal extraction from noise for a piece of time sequence

$$I(t) = \begin{cases} A\cos(\omega t + \alpha) + \Gamma(t), & 0 < t < T, \\ 0, & \text{otherwise}, \end{cases}$$
$$\langle \Gamma(t)\Gamma(t+\tau)\rangle = 2D\delta(\tau), \quad D = O(1), \tag{2}$$

where all D, A, ω (also $T_0 = 2\pi/\omega$), and T are supposed to be known. Assume a task that an active detector emits a sinusoidal signal with a given known frequency ω and time duration T, the signal is reflected (or not reflected) back from certain object in some unknown time interval. The amplitude of the reflected signal A can be well estimated for the detection task, and the noise background D can be easily computed from the received data. The central problem is greatly simplified to answer whether the sinusoidal signal with known amplitude A exists in a given received data of $I(t)$ (Eq. (2)). The most direct method to extract the signal[10] is to compute Fourier

spectrum at frequency ω

$$B=\sqrt{B_1^2+B_2^2},$$
$$B_1=\frac{2}{T}\int_0^T I(t)\cos(\omega t)\,\mathrm{d}t,$$
$$B_2=\frac{2}{T}\int_0^T I(t)\sin(\omega t)\,\mathrm{d}t. \tag{3}$$

For $T\to\infty$, we have $B=A$, and the signal can be definitely identified to exist or not from the computed B value (referring to Fig. 1(d) for large while still finite T). For small T, such as $T=50T_0$, $B\simeq A$ is no longer valid (referring to Fig. 1(c)). Nevertheless, with a certain value of the computed B, one can give an answer on probability only with which $I(t)$ includes the given signal. In the present work, all of our results rely on systematical probability computation that is important in realistic detections of signals with short durations, which has been seldom used so far, to the best of our knowledge, in previous works signal detections were performed often with infinitely long life time.

For any algebraic summations of Gaussian distributions must be a Gaussian distribution, we can generally write the (B_1, B_2) distribution as a two-dimensional joint Gaussian distribution

$$f(B_1,B_2)=\frac{1}{2\pi\sqrt{\sigma_1\sigma_2(1-\rho^2)}}\exp\left\{-\left[\frac{(B_1-\langle B_1\rangle)^2}{2\sigma_1}+\frac{(B_2-\langle B_2\rangle)^2}{2\sigma_2}-\rho\frac{(B_1-\langle B_1\rangle)(B_2-\langle B_2\rangle)}{\sqrt{\sigma_2\sigma_2}}\right]\frac{1}{1-\rho^2}\right\}. \tag{4}$$

Here $\langle B_1\rangle$ and $\langle B_2\rangle$ are the mean values of B_1 and B_2, σ_1 and σ_2 are the corresponding variances of B_1 and B_2, σ_{12} are the covariance of B_1 and B_2, and $\rho=\sigma_{12}/\sqrt{\sigma_1\sigma_2}$. By means of coordinate transformation from Cartesian (B_1, B_2) to polar (B,φ) $(B=\sqrt{B_1^2+B_2^2}, \tan\varphi=B_2/B_1)$, the distribution of the amplitude of the composition can be written as

$$P(B)=\frac{B}{2\pi\sqrt{\sigma_1\sigma_2(1-\rho^2)}}\cdot\int_0^{2\pi}\exp\left\{-\left[\frac{(B\cos\varphi-\langle B_1\rangle)^2}{2\sigma_1}+\frac{(B\sin\varphi-\langle B_2\rangle)^2}{2\sigma_2}-\rho\frac{(B\cos\varphi-\langle B_1\rangle)(B\sin\varphi-\langle B_2\rangle)}{\sqrt{\sigma_1\sigma_2}}\right]\frac{1}{1-\rho^2}\right\}\mathrm{d}\varphi. \tag{5}$$

For $I(t)=s(t)+\Gamma(t)$ with variables $s(t)$ and $\Gamma(t)$ being given in Eq. (2), the distribution of B can be specified as the following expression

$$P(B,A)=\frac{TB}{8\pi D}\exp\left[-\frac{T}{8D}(A^2+B^2)\right]\cdot\int_0^{2\pi}\exp\left(\frac{TBA}{4D}\cos\varphi\right)\mathrm{d}\varphi. \tag{6}$$

Without signal $A=0$, the formula can be further simplified to the following form

$$P(B,A=0)=\frac{BT}{4D}\exp\left(-\frac{T}{8D}B^2\right). \tag{7}$$

With the given parameters A, T, and D, the computed B can appear anywhere in the range of $(0,\infty)$ with different probabilities. For $A^2T/D\to\infty$, equation (6) reduces to the deterministic result: a sharp Gaussian probability distribution centered at $B=A$. Now the problem to detect whether a signal $A\cos(\omega t+\alpha)$ exists in $I(t)$ in a piece of received data of length T can be answered in a probabilistic way by comparing the probability distributions of the cases: Eq. (7) $A=0$ and Eq. (6) $A\neq 0$.

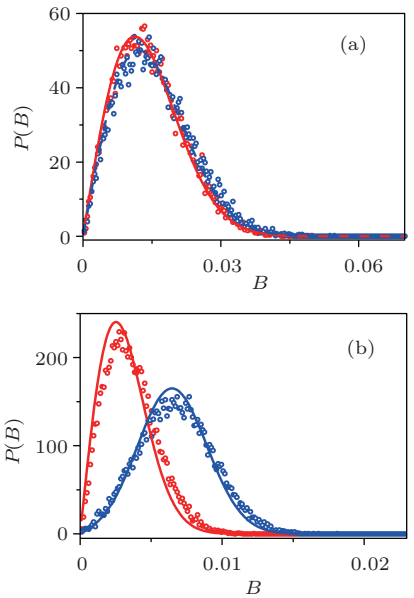


Fig. 2. (color online) Probability $P(B)$ computed by Eq. (6) (continuous curves) and by numerically realizations of Eq. (3) (dots) for the stochastic input $I(t)$ with ($A=0.006$, blue curve and dots) and without ($A=0$, red curve and dots) signal. The task is to answer whether a given received input data of length T includes signal $s(t)$. 2×10^4 realizations are used for averages of numerical dots. (a) $T=50T_0$. Both distributions $P(B)$'s of $A=0$ (red) and $A\neq0$ (blue) strongly overlap. It is difficult to make correct prediction. (b) The same as panel (a) with $T=100T_0$. With longer T the two distributions of $A=0$ and $A\neq0$ separated considerably. The probability to make correct conclusion becomes higher than panel (a). In both panels (a) and (b) theoretical predications of Eq. (6) and numerical results coincide with each other very well.

In Figs. 2(a) and 2(b) we fix different finite T's for no signal $A = 0$ (red curves) and with signal $0 < A^2/D \ll 1$ (blue curves), and plot the theoretical results (continuous curves) and numerical computations (dots), both results agree with each other perfectly.

3. Detection of weak and shortly-lived signal from strong noise by using resonant filters

Now we inject the received data $I(t)$ into a resonant filter[16]

$$\ddot{x} + \gamma\dot{x} + \omega^2 x = I(t). \tag{8}$$

According to Langevin theory, the transient solution of the system can be written as

$$x(t) = x_1(t) + x_2(t) + x_3(t), \tag{9}$$

where

$$\begin{aligned}
x_1(t) &= \eta \exp\left(-\frac{\gamma}{2}t\right)\sin(\omega_{\rm f}t) + x_0 \exp\left(-\frac{\gamma}{2}t\right)\cos(\omega_{\rm f}t),\\
x_2(t) &= \frac{A'}{\gamma\omega}\sin(\omega t + \alpha),\\
x_3(t) &= \exp\left(-\frac{\gamma}{2}t\right)[R_2(t)\sin(\omega_{\rm f}t) - R_1(t)\cos(\omega_{\rm f}t)],
\end{aligned}$$

with

$$\begin{aligned}
R_1(t) &= \frac{1}{\omega_{\rm f}}\int_0^t \exp\left(\frac{\gamma}{2}s\right)\sin(\omega_{\rm f}s)\Gamma(s)\,{\rm d}s,\\
R_2(t) &= \frac{1}{\omega_{\rm f}}\int_0^t \exp\left(\frac{\gamma}{2}s\right)\cos(\omega_{\rm f}s)\Gamma(s)\,{\rm d}s,\\
\omega_{\rm f} &= \sqrt{\omega^2 - \frac{\gamma^2}{4}}, \quad A' = \sqrt{x_0^2 + \eta^2},\\
\eta &= \frac{(\dot{x}_0 + \gamma x_0/2 - A/\gamma)}{\omega_{\rm f}}.
\end{aligned}$$

In the above expressions, $(x_0, \dot{x}_0)$ is the initial condition, $A\cos(\omega t + \alpha) + \Gamma(t)$ is the stochastic input with signal, and γ is the damping coefficient of the linear filter system. The white noise, in terms of stochastic differential equations, corresponds to the Stratonovich formalism.[16]

Since this equation is exactly solvable, it seems that there is nothing interesting with further analysis. For $T \to \infty$ for any initial condition $(x_0, \dot{x}_0)$ the system evolves to an asymptotic destination where the initial condition is forgotten, and SNR of $x(t)$ is exactly the same as SNR of $I(t)$ due to the linearity of the dynamics. However, for T being finite, the case turns to be much more complicated. The system is still solvable, but the solution turns to be strongly nonlinear with the initial condition $(x_0, \dot{x}_0)$ where noise, signal, and all system parameters interact in a very complex manner. Some interesting results appear by these interactions.

The signal can be extracted from $x(t)$ for finite time sequences of length NT_0. By numerical computations, similar to that for $I(t)$, here B is given by

$$\begin{aligned}
B &= \sqrt{B_1^2 + B_2^2},\\
B_1 &= \frac{2}{T}\int_0^T x(t)\cos(\omega t)\,{\rm d}t,\\
B_2 &= \frac{2}{T}\int_0^T x(t)\sin(\omega t)\,{\rm d}t.
\end{aligned} \tag{10}$$

The computational realization of the distributions of $P(B, A \neq 0)$ and $P(B, A = 0)$ are dotted in Fig. 4, comparing to the inputs in Fig. 2, the two distribution curves seemly separate more from each other, and thus the filter makes some difference indeed.

Probability of $P(B)$ can be theoretically carried out, similar to that for the input, though the former is much more tedious than the latter and the results are not explicit. Analytic and exact formula of Eq. (5) for $x(t)$ with different mean values and variances can be reached. From the deterministic part of Eq. (9), $x_1(t)$ and $x_2(t)$, the mean values of B_1 and B_2 can be obtained as

$$\begin{aligned}
\langle B_1\rangle &= \frac{2x_0}{NT_0\gamma}\left[1 - \exp\left(-\frac{\gamma}{2}NT_0\right)\cos(\omega_{\rm f}NT_0)\right] - \frac{2\eta}{NT_0\gamma}\exp\left(-\frac{\gamma}{2}NT_0\right)\sin(\omega_{\rm f}NT_0),\\
\langle B_2\rangle &= \frac{1}{NT_0\omega}\left\{\left(x_0\frac{2\omega_{\rm f}}{\gamma} - \eta\right)\exp\left(-\frac{\gamma}{2}NT_0\right)\sin(\omega_{\rm f}NT_0) + \left(x_0 + \eta\frac{2\omega_{\rm f}}{\gamma}\right)\left[1 - \exp\left(-\frac{\gamma}{2}NT_0\right)\cos(\omega_{\rm f}NT_0)\right]\right\} + \frac{A}{\gamma\omega}.
\end{aligned} \tag{11}$$

The variances of B_1, B_2, and their covariance define σ_1, σ_2, and σ_{12}. All of them depend on the random item $x_3(t)$ only, and they read

$$\sigma_j = \frac{8}{N^2T_0^2}\int_0^{NT_0}\int_0^{t_2}\langle x_3(t_1)x_3(t_2)\rangle h_j\,{\rm d}t_1\,{\rm d}t_2, \quad j = 1, 2, 12, \tag{12}$$

with $h_1 = \cos\omega t_1\cos\omega t_2$, $h_2 = \sin\omega t_1\sin\omega t_2$, and $h_{12} = \cos\omega t_1\sin\omega t_2$, where

$$\langle x_3(t_1)x_3(t_2)\rangle = \frac{2D}{\omega_{\rm f}^2}\exp\left[-\frac{\gamma(t_1 + t_2)}{2}\right]$$

$$\cdot\Big\{\sin(\omega_{\rm f}t_1)\sin(\omega_{\rm f}t_2)\int_0^{t_2}\int_0^{t_1}\exp\Big[\frac{\gamma}{2}(x+y)\Big]\cos(\omega_{\rm f}x)\cos(\omega_{\rm f}y)\delta(x-y)\,{\rm d}x\,{\rm d}y$$
$$-\cos(\omega_{\rm f}t_1)\sin(\omega_{\rm f}t_2)\int_0^{t_2}\int_0^{t_1}\exp\Big[\frac{\gamma}{2}(x+y)\Big]\sin(\omega_{\rm f}x)\cos(\omega_{\rm f}y)\delta(x-y)\,{\rm d}x\,{\rm d}y$$
$$-\sin(\omega_{\rm f}t_1)\cos(\omega_{\rm f}t_2)\int_0^{t_2}\int_0^{t_1}\exp\Big[\frac{\gamma}{2}(x+y)\Big]\cos(\omega_{\rm f}x)\sin(\omega_{\rm f}y)\delta(x-y)\,{\rm d}x\,{\rm d}y$$
$$+\cos(\omega_{\rm f}t_1)\cos(\omega_{\rm f}t_2)\int_0^{t_2}\int_0^{t_1}\exp\Big[\frac{\gamma}{2}(x+y)\Big]\sin(\omega_{\rm f}x)\sin(\omega_{\rm f}y)\delta(x-y)\,{\rm d}x\,{\rm d}y\Big\}. \tag{13}$$

So that, equations (12) can be expressed into a unified form with

$$\sigma_j = g_0\Big\{g_{j,1}NT_0 + g_{j,2}\Big[1-\exp\Big(-\gamma NT_0\Big)\Big] + g_{j,3}\exp\Big(-\frac{\gamma}{2}NT_0\Big)\sin(\omega_{\rm f}NT_0)$$
$$+\,g_{j,4}\Big[1-\exp\Big(-\frac{\gamma}{2}NT_0\Big)\cos(\omega_{\rm f}NT_0)\Big] + g_{j,5}\exp(-\gamma NT_0)\sin(2\omega_{\rm f}NT_0) + g_{j,6}\big[1-\exp(-\gamma NT_0)\cos(2\omega_{\rm f}NT_0)\big]\Big\},$$
$$g_0 = \frac{1}{N^2T_0^2}\frac{D}{\omega\omega_{\rm f}^2},\quad j=1,2,12,$$
$$g_{1,1}=\frac{4\omega_{\rm f}^2}{\omega\gamma^2},\quad g_{1,2}=\frac{4\omega}{\gamma^3},\quad g_{1,3}=\frac{8\omega_{\rm f}}{\omega\gamma^2},\quad g_{1,4}=-\frac{16\omega_{\rm f}^2}{\omega\gamma^3},\quad g_{1,5}=-\frac{2\omega_{\rm f}}{\omega\gamma^2},\quad g_{1,6}=-\frac{1}{\omega\gamma},$$
$$g_{2,1}=\frac{4\omega_{\rm f}^2}{\omega\gamma^2},\quad g_{2,2}=\frac{4\omega}{\gamma^3},\quad g_{2,3}=\frac{8\omega_{\rm f}}{\omega\gamma^2},\quad g_{2,4}=-\frac{16\omega_{\rm f}^2}{\omega\gamma^3},\quad g_{2,5}=\frac{\omega_{\rm f}}{\omega}\Big(\frac{2}{\gamma^2}-\frac{2}{\omega^2}\Big),\quad g_{2,6}=\frac{\gamma}{\omega}\Big(\frac{3}{\gamma^2}-\frac{1}{\omega^2}\Big),$$
$$g_{12,1}=-\frac{4\omega_{\rm f}^2}{\omega^2\gamma},\quad g_{12,2}=\frac{4}{\gamma^2},\quad g_{12,3}=\frac{16\omega_{\rm f}}{\gamma^3},\quad g_{12,4}=0,\quad g_{12,5}=-\frac{\gamma\omega_{\rm f}}{\omega^4}-\frac{\omega_{\rm f}}{\omega^2\gamma},\quad g_{12,6}=\frac{1}{2\omega^2}-\frac{\gamma^2}{2\omega^4}+\frac{2}{\gamma^2}. \tag{14}$$

Finally, probability $P(B)$ of $x(t)$ for finite time sequences of length NT_0 can be solved analytically while implicitly by inserting the solutions (11) and (14) to Eq. (5).

In Eqs. (10)–(14), all the controllable matters of the filter, including parameters γ [related to the quality factor of the filter, which also take effect through $\omega_{\rm f}$ referring to Eq. (9)], and initial values of $(x_0,\dot{x}_0)$ interact with the facts of the input NT_0 and ω in strongly nonlinear manner, and effectively influence the probability distribution. The central task of the present paper is schematically shown in Fig. 3, how the above facts alter the resulting probability distributions and whether the efficiency of the weak signal detection by the resonant filter can out-perform the algorism of Eq. (5) for the input.

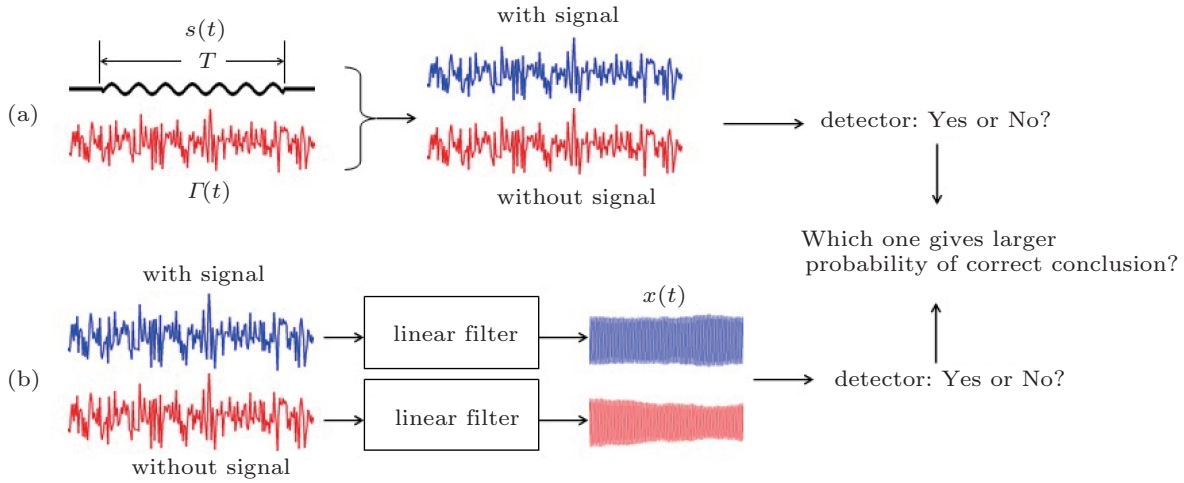


Fig. 3. (color online) Schematic figures of signal detection tasks. (a) Direct signal detection from the input $I(t)$. (b) Signal detection from the output of a selected linear filter $x(t)$ driven by the input $I(t)$. The central problem is: whether the latter analysis can out-perform the former signal detection directly from the input.

With the two different inputs $I(t)$ of Figs. 2(a) and 2(b) we compute, in Figs. 4(a) and 4(b), data of output $x(t)$'s from filter (8) for certain optimal parameters. A number of features of the results are interesting. First, the theoretical predictions of Eqs. (5), (11), and (14) coincide with numerical plots very well. Second, we find interestingly that data for $x(t)$ in Figs. 4(a) and 4(b) are seemly better than those for $I(t)$ in Figs. 2(a) and 2(b) by providing larger separations of the two compared distributions. In particular, for rather short T, the distributions of A=0 and $A\neq 0$ in Fig. 4(a) are separated

more than those in Fig. 2(a).

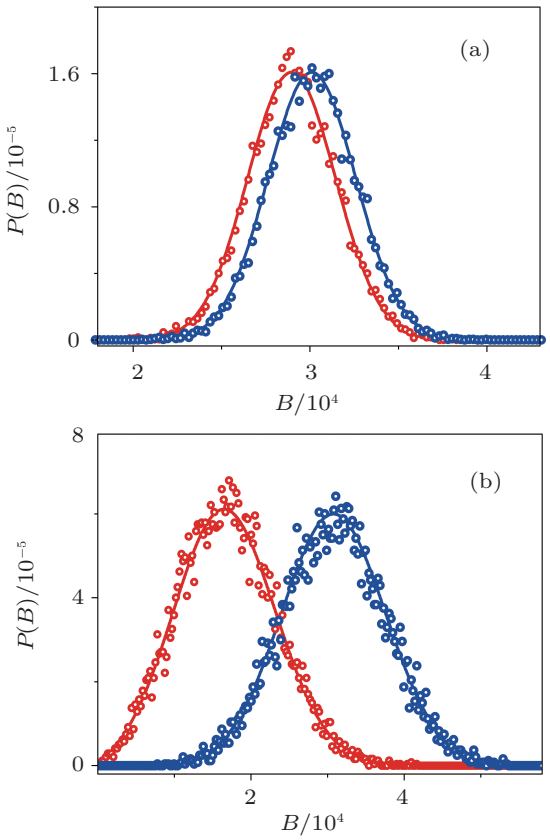


Fig. 4. (color online) The same as Fig.2 with formula of Eqs. (5), (11), and (14) (continuous curves) and numerical results of Eq. (10) (dots) applied. Parameters of the filter Eq. (8) are given as: $\gamma = 10^{-5}$, $x_0 = 0$, and $\dot{x}_0 = A/\gamma$. The input $I(t)$ in Fig. 2 is replaced by the output $x(t)$ of Eq. (8).

4. How to enhance the efficiency of detections of weak and shortly durative signals by adjusting parameters of resonant filters

It is emphasized that with available known information we can well draw the two distributions while we cannot control the value of B for each realization, and cannot definitely conclude which B corresponds signal (or no signal). The most probabilistic way to adjudge whether $I(t)$ includes ($A \neq 0$) or not includes ($A = 0$) signal with Eq. (6) with a given computed B is: Yes (blue) for $P(B,A \neq 0) > P(B,A = 0)$ or no (red) for $P(B,A \neq 0) \leqslant P(B,A = 0)$. For quantitatively measuring the efficiency of signal extraction with the computed B values in Eqs. (3) and (10), we define a quantity of separation of probability distributions of data with and without signal, based on the measurement of various areas of Fig. 5(a). Taking the two curves in Fig. 2(b) as examples, the two distributions have equal probability density at $B_C \approx 0.00473$ (dotted line in Fig. 5(a)). We have $P(B,A \neq 0) \gtrless P(B,A = 0)$ for $B \gtrless B_C$, respectively. Therefore, we more likely estimate $A > 0$ (signal received) for any $B > B_C$ while $A = 0$ (no signal received) for $B < B_C$. The correctness percentage for these probabilistic estimations read

$$P_S = \frac{M_1}{M_1 + M_4}, \quad P_N = \frac{M_3}{M_2 + M_3}. \tag{15}$$

Here P_S is the probability of successful signal extractions and P_N is the probability of successful identification of no signal. In both cases, wrong estimations come from the overlapped area of the two distributions (with and without signal). We can thus simply define the Separation of Probability Distributions (SPD) of the two cases

$$\mathrm{SPD} = \frac{M_1 + M_3}{2} \tag{16}$$

as the measure of the overall probability of correct signal (or no signal) detection. Larger SPD represents better signal extraction.

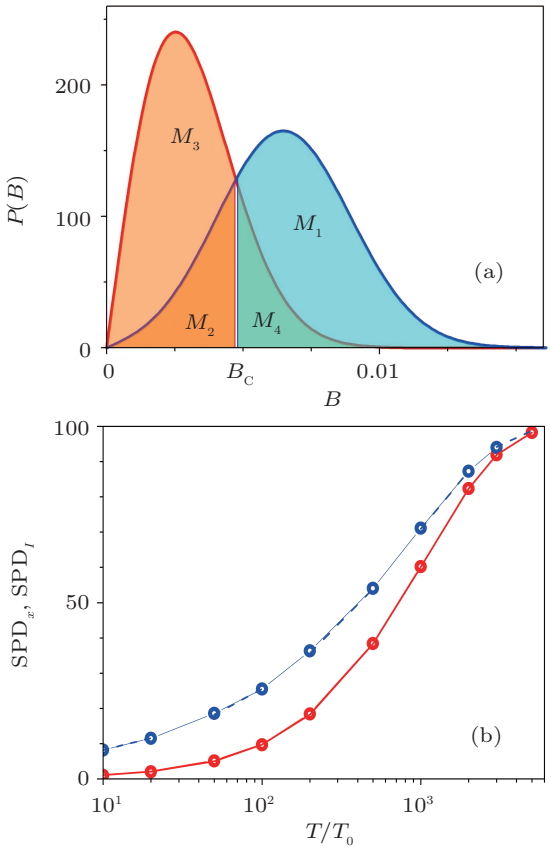


Fig. 5. (color online) Definition and computation of separation of probability distributions SPD with signal from that without signal. (a) Schematic figure on signal prediction and definition of SPD. At $B = B_C$ we have $P(B_C,A \neq 0) = P(B_C,A = 0)$, and $P(B_C,A \neq 0) \gtrless P(B_C,A = 0)$ for $B \gtrless B_C$, respectively. We thus most probably predict signal when $B > B_C$ while no signal for $B < B_C$. And the overall probability of correct predictions (including correct predictions of signal and no signal) reads $\mathrm{SPD} = (M_1 + M_3)/2$. (b) SPD_I for the input $I(t)$ of Eqs. (3) and (6) (red colored plots) and SPD_x for the output $x(t)$ of Eqs. (10) and (5), (11), (14) (blue colored plots) plotted against T. $\mathrm{SPD}_x > \mathrm{SPD}_I$ is identified by theoretical and rigorous derivations (continuous curves) and numerical computations (dots). In particular, ratio $\mathrm{SPD}_x/\mathrm{SPD}_I$ is rather larger than 1 for small T's.

Equation (16) can be analytically computed for both input of $I(t)$ (Eq. (2)) and output $x(t)$ (Eq. (8)), and can be also numerically computed, based on various noise realizations. In

Fig. 5(b) we plot SPD's for $I(t)$ and $x(t)$ against different T's with other parameters fixed as $\gamma = 10^{-5}$, $x_0 = 0$, and $\dot{x}_0 = A/\gamma$. A striking observation is that the SPD values of output $x(t)$ are considerably higher than those of input $I(t)$. And for small T's the former is more than several times higher than the latter. This is the very first time that a rigorous analysis shows possibility to enhance the efficiency of signal detection. All the theoretical results (solid lines) are perfectly confirmed by numerical computations (dots).

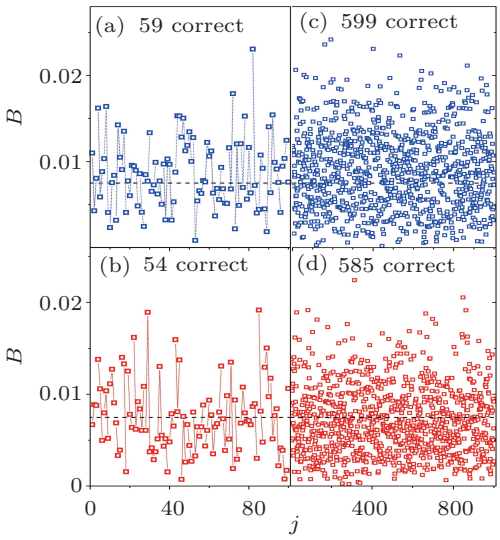


Fig. 6. (color online) Signal predictions of 100 [(a) and (b)] and 1000 [(c) and (d)] realizations by numerically computing Eq. (3) for $I(t)$ without signal ($A = 0$, (a) and (c)) and with signal ($A = 0.006$, (b) and (d)). $D = 0.5$, $T = 200T_0$. Numbers of correct predictions are written in the corresponding legends.

In Fig. 6 we show correct and incorrect conclusions of signal detection for $I(t)$'s with 200 and 2000 realizations, 100 and 1000 for $A = 0$, and 100 and 1000 for $A \neq 0$. In Fig. 7 we do exactly the same as Fig. 6 by using the same $I(t)$'s, but with the outputs of the filter Eq. (8), $x(t)$ data, considered. By choosing proper parameters we obtain better detection results, for obviously more realizations move to their correct areas than the realizations oppositely moved.

The resonant filter does improve the signal detection referring to Fig. 8, where we do the same as Fig. 6 (panels (c) and (d)) and Fig. 7 (panels (c) and (d)) while by ten times. It is striking that, each time the movements of plots from $I(t)$ to $x(t)$ are in overall towards to directions of considerable improvements. And the improvements have rather high ratios for short T's as shown in Table 1.

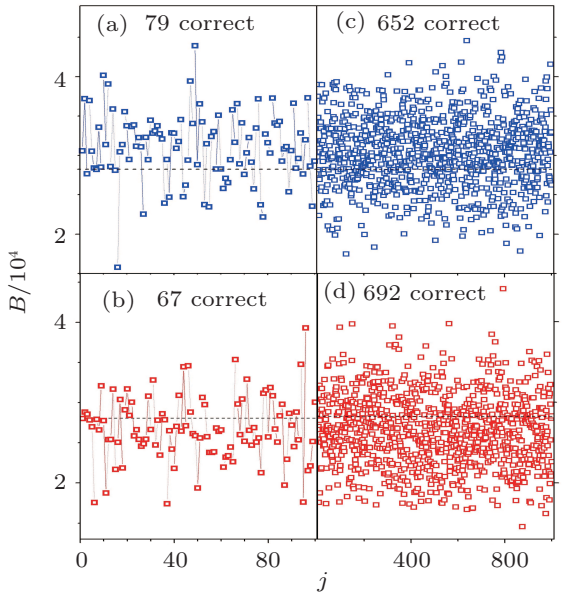


Fig. 7. (color online) (a)–(d) The same as Figs. 6(a)–6(d), respectively, by numerically computing Eq. (10) for $x(t)$. Parameters of the resonant filter are: $\gamma = 10^{-5}$, $x_0 = 0$, and $\dot{x}_0 = A/\gamma$. In each case the detections with $x(t)$ here have more numbers of correct conclusions than those with $I(t)$ in Fig. 6.

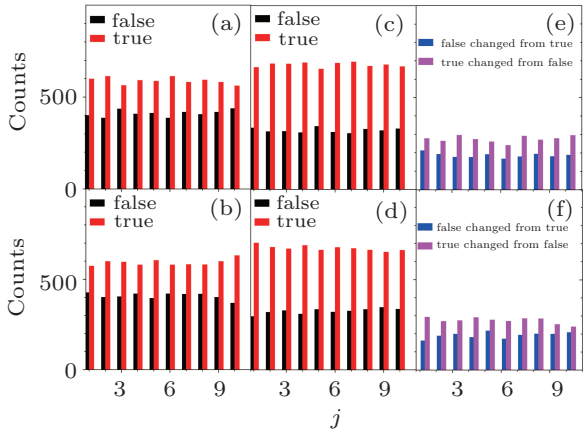


Fig. 8. (color online) Bar charts of signal detection results on ten groups of experiments of Figs. 6(c) and 6(d) and Figs. 7(c) and 7(d). Each group performs 1000 realizations of numerical computations of Eq. (3) (the input: (a) for $A = 0$, (b) for $A = 0.006$) and Eq. (10) (the output: (c) for $A = 0$, (d) for $A = 0.006$). Numbers of correct (red) and error (black) detections are ranged. It is obvious that with the optimal resonant filter more realizations move to the correct side than those moving to the wrong side (panels (e) and (f)).

Table 1. Dependence of SPD's of $I(t)$ and $x(t)$, and their ratios on the time length.

	$10T_0$	$20T_0$	$50T_0$	$100T_0$	$200T_0$	$500T_0$	$1000T_0$	$2000T_0$	$3000T_0$
SPD_I	1.07	2.05	5.02	9.71	18.21	38.26	60.09	82.33	91.91
	(1.12)	(2.09)	(5.15)	(9.78)	(18.43)	(38.41)	(60.02)	(82.36)	(91.96)
SPD_x	8.19	11.59	18.24	25.60	35.67	53.96	70.92	87.22	94.14
	(8.23)	(11.47)	(18.56)	(25.52)	(36.32)	(54.02)	(71.11)	(87.22)	(94.01)
$\mathrm{SPD}_x/\mathrm{SPD}_I$	7.65	5.65	3.63	2.64	1.96	1.41	1.18	1.06	1.02
	(7.35)	(5.49)	(3.60)	(2.61)	(1.97)	(1.41)	(1.18)	(1.06)	(1.02)

SPD's are computed by theoretical formulas (6) and (5), (11), (14) (without bracket) and numerical simulations (in brackets). $A = 0.006$, $D = 0.5$. Initial conditions for the linear filter are $x_0 = 0$, and $\dot{x}_0 = A/\gamma$, and other parameters $\gamma = 10^{-5}$, $\omega = 0.02$ (also $T_0 = 2\pi/\omega$) are taken.

Various parameters such as γ, x_0, and $\dot{x}_0$ are included in $P(B)$ distribution of Eqs. (5), (11), and (14). It is interesting to study how these additional physical quantities in the filter, beyond the input data $I(t)$, influence the detection efficiency.

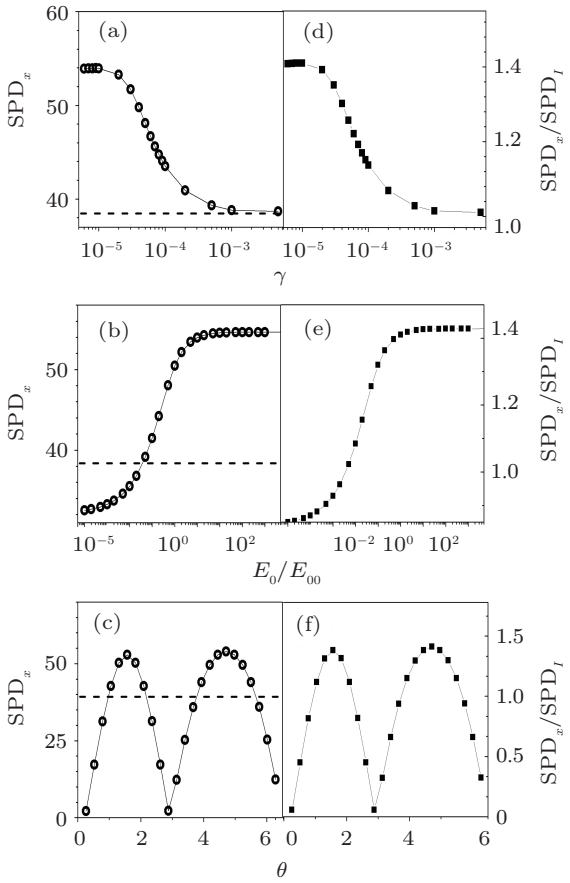


Fig. 9. Dependence of signal detection efficiency on control parameters. $D = 0.5$, $A = 0.006$, $\omega = 0.02$, $T = 500T_0$. (a) SPD$_x$ plotted against γ with $E_0 = A^2/(2\gamma^2)$ and $\theta_0 = -\pi/2$. SPD$_I = 38.26$ for $I(t)$ is drawn for comparisons. There is an optimal $\gamma \approx 10^{-5}$, providing the largest SPD$_x$ which is much higher than SPD$_I$. (b) The same as panel (a) with SPD$_x$ plotted versus E_0 with $\gamma = 10^{-5}$ and $E_{00} = A^2/(2\gamma^2)$. SPD$_x$ increases as E_0 and saturates to a value SPD$_x >$ SPD$_I$. (c) The same as panel (a) with SPD$_x$ versus θ for $\gamma = 10^{-5}$. SPD$_x$ reaches its maximum at $\theta_0 = \pm\pi/2$. The results of panels (a), (b), and (c) show convincingly that the additional parameters of the filter can dramatically influence the efficiency of signal detection, and enhance the efficiency considerably higher than the input at some optimal parameter sets. Panels (d), (e), and (f) represent ratios SPD$_x$/SPD$_I$ corresponding to panels (a), (b), and (c), respectively.

In Fig. 9(a) we plot SPD$_x$ against γ with other parameters fixed, the probability of correct signal detection increases as γ reduces, and approaches saturation around $\gamma \approx 10^{-5}$. In Fig. 9(b) we plot SPD$_x$ versus initial energy preparation E ($E = (\omega^2 x^2 + \dot{x}^2)/2$) of the filter, SPD$_x$ increases monotonously as E and saturates to a certain value higher again than that of the input. In Fig. 9(c) SPD$_x$ is plotted against the initial phase angle θ (ctg $\theta = x_0\omega/\dot{x}_0$). There exists certain optimal phases, where SPD of the output $x(t)$ is considerably better than that of the input $I(t)$. All plots in Fig. 9 and Fig. 10 show convincingly that the controllable parameters in the filter detector are effectively involved in the signal extractions, and optimal match between these parameters can indeed do better in performing this task. Roughly speaking, smaller γ, larger E, and optimal initial phase arrangement can provide better results of signal detections. It is again emphasized that all numerical simulations well agree with the theoretical statistical analyses of Eqs. (6) and (5), (11), (14).

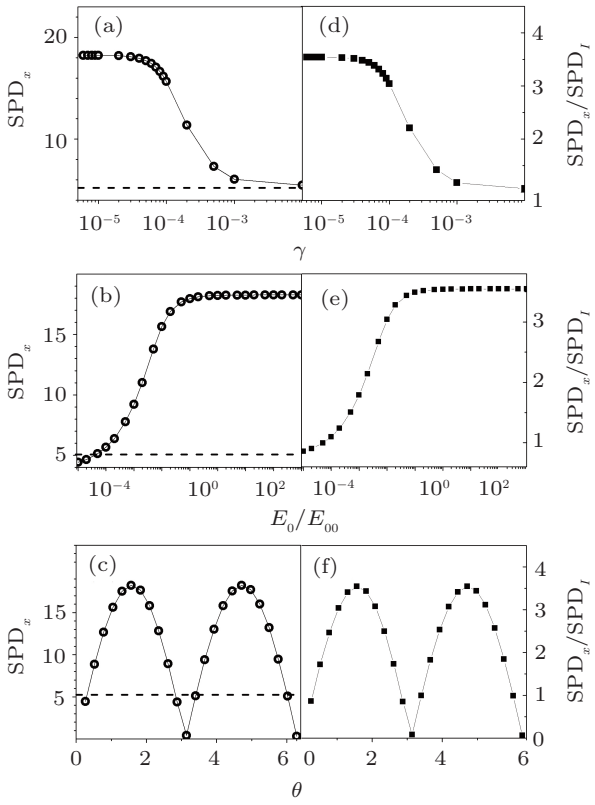


Fig. 10. The same as Fig. 9 except $T = 50T_0$.

5. Conclusion

In conclusion, we have studied the problem of signal detection by focusing on the simplest and also most fundamental cases of white-noise-contaminated sinusoidal signal. The main characteristics of the task are weak signal, strong noise and finite length of received data. In these cases, the signal detection can be predicted only by probabilistic description. Detectors of linear resonant filters are designed to enhance the efficiency of signal detections. While for infinite length of input data, the results of signal detection is irrelevant of all parameters of detectors which cannot improve the detection efficiency due to the linearity of the detector dynamics, for finite time signal length all these parameters, including the initial preparations of detector states, can effectively change the detection results by interacting with the input signal and noise in highly nonlinear manner. At certain optimal parameter matches the designed method can improve the signal detection efficiency much better than that of the input. All the above results are derived analytically and rigorously, and numerical computa-

tions fully support the theoretical predictions.

In this work our analyses are based on much simplified situation of noisy sinusoidal signals with known ω, T, D, and even A, and the problem is only to answer: yes or no. Nevertheless, even in these simplest cases, the problem whether we can improve signal detection, out-performing the standard algorism of input B remains a problem with long debates and conflicts, and there has been so far no any rigorous analysis clearly answering this problem. Here, we give this fundamental problem a definite and positive answer.

Although we focused on extracting sinusoidal signal from strong noise contamination, this is the simplest as well as the fundamental signal detection problem. The extension to more general and also more difficult problems of extracting general periodic or non-periodic signals from noise background is of practical significance. We expect to go further to attack these goals in our future works.

With various deeper analyses and more effective improvements by considering more realistic facts, the results are expected to be applicable to practical signal detection problems where weak signal, strong noise and limited available data length are crucial.

References

[1] Green D M 1958 *J. Acoust. Soc. Am.* **30** 904
[2] Martin R D and Schwartz S C 1971 *IEEE Trans. Inform. Theory* **17** 50
[3] Wang Y, Xu G, Liang L and Jiang K 2015 *Mechanical Systems and Signal Processing* **54–55** 259
[4] Krasnenker V M 1980 *Automation and Remote Control* **41** 640
[5] Jung P and Hänggi P 1991 *Phys. Rev. A* **44** 8032
[6] Inchiosa M E and Bulsara A R 1996 *Phys. Rev. E* **53** R2021
[7] Duan F, Chapeau-Blondeau F and Abbott D 2013 *Digital Signal Processing* **23** 1585
[8] Siddagangaiah S, Li Y, Guo X, Chen X, Zhang Q, Yang K and Yang Y 2016 *Entropy* **18** 101
[9] Hänggi P, Inchiosa M, Fogliatti D and Bulsara A R 2000 *Phys. Rev. E* **62** 6155
[10] Casado-Pascual J, Denk C, Gómez-Ordóñez J, Morillo M and Hänggi P 2003 *Phys. Rev. E* **67** 036109
[11] Jimenez-Aquino J and Romero-Bastida M 2011 *Phys. Rev. E* **84** 011137
[12] Duan F, Chapeau-Blondeau F and Abbott D 2014 *Phys. Rev. E* **90** 022134
[13] Khovanov I A 2008 *Phys. Rev. E* **77** 011124
[14] Ou Z Y 2012 *Phys. Rev. A* **85** 023815
[15] He L F, Cui Y Y, Zhang T Q, Zhang G and Song Y 2016 *Chin. Phys. B* **25** 060501
[16] Risken H 1984 *The Fokker–Planck Equation* (Berlin: Springer-Verlag) pp. 63–110, pp. 229–275