# An adaptive stochastic-resonance-based detector and its application in watermark extraction

Gencheng Guo and Mrinal Mandal
Department of Electrical and Computer Engineering
University of Alberta
9107 - 116 Street, Edmonton
CANADA, T6G 2V4
gencheng@ece.ualberta.ca, mandal@ece.ualberta.ca, http:// www.ece.ualberta.ca/ mandal/

*Abstract:* - In this paper, we explore a stochastic resonance (SR) based detector using bistable system (BS) to detect a binary pulse amplitude modulated (PAM) signal embedded in non-Gaussian noise. Through the example of BS based watermark extraction, we show that a reliable performance cannot be obtained if the BS parameters are determined by traditional tuning technique. The key observation is that the BS parameters are not sensitive to the pdf of the noise but to the variance of the noise and the amplitude of the signal. That makes it possible to determine the BS parameters in advance and an adaptive BS can be constructed based on the estimated amplitude of the watermark (signal) and the variance of the DCT coefficients (noise). Experimental results show that the performance obtained from the proposed adaptive stochastic-resonator-based detector is stable and provides superior performance compared to the existing BS based watermark schemes and the Gaussian based maximum likelihood (ML) detector.

*Key-Words:* - Stochastic resonance, bistable system, optimal parameters, watermark extraction.

## 1 Introduction

Stochastic resonance (SR) phenomenon was proposed by Benzi et al. to model the periodically recurrent ice ages [1]. Since then the SR has been used in many engineering systems to perform a wide variety of tasks [2], such as sensory enhancement and signal detection. The signature of SR is that the coherence of a system output improves with an increase of the noise. The output performance is revealed as a non-monotonic evolution that increases to a peak value with the increase of the noise and then decreases with further increase of the noise, which is typically termed as SR effect. In this case, the additional noise is added to the fixed system and the SR effect occurs. Therefore, the phenomenon is referred to noise-enhanced SR (NSR). On the other hand, parameter-induced SR (PSR) describes the phenomenon that the input is fixed and system parameters are tuned to have a SR effect [3]-[5]. Both NSR and PSR have been employed in signal and image processing and have been shown to provide an improvement in weak signal enhancement [3]-[10].

SR has been used in watermark extraction because of the characteristic of enhancement of the weak signal [11]-[14]. In the SR based digital watermarking techniques, watermark information considered as signal embeds in discrete cosine transform (DCT) coefficients regarded as noise. Because the signal and noise are fixed, the system parameters are adjusted to obtain the optimal performance of the extraction. Sun et al. proposed to use the SR effect in watermark extraction [11],[12]. Here, the watermark information is added to the permuted DCT coefficients of the host image, and a bistable system (BS) is used to extract the watermark. Wu et al. [13] employed the same strategy but provided some improvement, such as using $8\times8$ DCT blocks and an effective permutation of DCT coefficients. Duan et al. [14] proposed to add a desynchronization delay to improve the robustness of the technique compared to the matched filter. This technique also uses an array of SR detectors to improve the performance. The BS is used in [11]-[14] to extract the watermarks and these extraction techniques can be considered as PSR based watermark extraction technique.

Although the PSR techniques mentioned above provide a good performance for watermark extraction, these techniques have a few limitations. These techniques typically decide one set of the BS parameters and use it to extract the watermarks from different images. It is observed that they use only one parameter set in [11]-[13] but the parameter sets are not identical. The choice of the SR parameters is critical in obtaining a good extraction performance. Hence we may ask the following questions. To extract watermark from varying watermarked images, can one parameter set always provide good performance of extraction? If yes, how can we determine this parameter set? If not, can we build a system adaptive to the varying images? Duan et al. proposed a tuning method to obtain the appropriate parameters [14], [15]. However, it is not feasible to find the optimal parameters by tuning when we extract the watermarks from an image due to the difficulties in estimating the extraction performance. The watermark extraction performance can be evaluated using objective criteria (e.g. BER) or subjective criteria (e.g. human eyes). However, for watermark technique, we usually prefer a blind watermark, and hence the objective criteria cannot be applied. For the subjective criteria, the performance could not be guaranteed and may not be used in many applications.

More generally, the problem of the watermark extraction is a problem of signal detection. Because the BS is widely used in exploiting the SR phenomena, the technique of the BS based watermark extraction is named as BS-SR watermark extraction [11]-[14]. Though not accurately named (we will show it below), the term is still used in this paper and henceforth the detector is referred to as the BS-SR detector.

The schematic of the BS-SR detector is shown in Fig. 1. Note that the BS-SR detector is constructed using a BS followed by an **inner detector** [6], [16]-[18] and BS shall be used as a stochastic resonator (or working in the SR regime). However, this is not the case in watermark extraction. As pointed out in [6], [19], conventional SR possesses the feature that noise plays a constructive role (for example, can increase the SNR), and therefore only happens when the signal is subthreshold or slightly suprathreshold in BS. While the BS employed in [11]-[14] is highly suprathreshold considering the amplitude of the signal (watermark embedded).

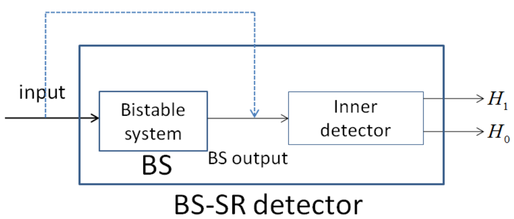



Fig. 1 The schematic of the BS-SR detector.

We have demonstrated in [20] that the BS is changed to a nonlinear component in a conventional detector if the BS parameters are obtained by tuning technique relying on the overall performance of the BS-SR detector. For the case of watermark extraction, it is shown that the probability density function (pdf) of the DCT coefficients are bell-shaped with heavy tails [21]. This kind of non-Gaussian noise exhibits spikes and a good detector typically includes a nonlinear limiter to reduce the noise spikes [22]. The BS can provide suitable nonlinearity to be a limiter and the BS-SR detector will change to a conventional detector, such as locally optimal (LO) detector [23]. Therefore, the BS employed in watermark extraction [11]-[14] could be considered as a limiter but not a stochastic resonator because tuning the BS parameters for a good overall detection performance will change the BS to a limiter that is the optimal nonlinearity in this case. The problem is that the tuning technique is not practical as mentioned in the drawbacks of the PSR base watermark extraction previously.

To clarify the function of the BS in the BS-SR detector, we proposed in [20] the design of the BS-SR detector by investigating the BS and inner detector separately. The BS parameters are determined according to a SR measure (cross-correlation) to make the BS work as a stochastic resonator and then design the inner detector depending on the statistic feature of the BS output. Note that here the BS used as stochastic resonator may not be subthreshold and therefore the BS is not a strict stochastic resonator (increasing noise will not lead to an increase of the SR measure). Here, however, the BS is investigated solely (hence will not be the limiter) and the BS parameters are chosen according to the SR measure. Therefore we still use BS-SR detector henceforth. The major benefit obtained from this set up is that the BS parameters can be determined in a systematic way.

In this paper, we address the design of the BS used as a stochastic resonator but not include the design of the inner detector that is assumed to be a linear detector. We have two major contributions. First, we propose a method to determine the BS parameters according to the SR measure. We investigate the working of the BS and show how the BS parameters are related to the SR measure. We then propose a tuning one parameter (TOP) technique to choose the near-optimal parameter sets to reduce the complexity of tuning. Furthermore, by analyzing the mechanism of the BS, in the view point of the SR measure, we find that the BS parameters are not sensitive to the pdf of the noise but only to the variance of the noise and the amplitude of the signal. Hence the BS parameters can be determined by an experimental way. Secondly, we go one step further to design a BS-SR detector adaptive to the varying input. It is illustrated that we cannot achieve a good performance by applying only one parameter set for varying images and hence the BS-SR detector with adaptive parameters is proposed. The Experimental results show that the performance obtained from the proposed adaptive BS-SR detector is stable and is superior compared to the existing BS based watermark schemes and the Gaussian based maximum likelihood (ML) detector.

The organization of this paper is as follows. In Section 2, we present a brief review of the BS-SR based watermarking techniques. In Section 3, we present the TOP technique to tune the BS and the method to determine the BS parameters. In Section 4, the BS-SR based watermark technique is analyzed and the adaptive BS-SR detector is proposed. We present the simulation results in Section 5, followed by conclusions and discussions in Section 6.

## 2 Review of the background work

In this section, we present a brief review of the background work. First, we introduce the BS used in the BS-SR based watermark extraction technique. We then present a review of the existing BS-SR based watermarking techniques.

### 2.1 Bistable system

A commonly used BS can be understood using an analogy with a double well potential shown in Fig. 2 and expressed by the following equation:

$$U(p) = -\left(\frac{a}{2}\right)p^2 + \left(\frac{b}{4}\right)p^4 \qquad (1)$$

where $(a,b)$ are parameters of the double well potential. The barrier height, i.e., the distance between central barrier and the well bottom, is $H(0) = a^2/b$ and potential minima are located at $\pm c = \pm\sqrt{a/b}$ . The bistable system is the evolvement of a heavily damped ball put in the double well. The system output is $p(t)$ , i.e., the position of the ball along $p$ axis. In the absence of any input, the velocity of the ball along $p$ axis is expressed by $\dot{p}(t) = ap(t) - bp^3(t)$ and the ball will fall down and rest in its equilibrium

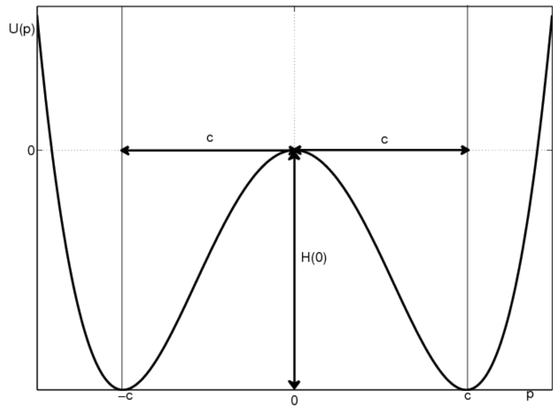



Fig. 2 Double well potential

point located at $\pm c$ . An input is the velocity added to the ball. Therefore, with input $x(t)$ the velocity of the ball will be

$$\dot{p}(t) = ap(t) - bp^3(t) + x(t). \qquad (2)$$

Because the closed-form solution of $p(t)$ is unavailable, we need to obtain $p(t)$ numerically. It is worth pointing out the discrete computer simulation model of Eq. (2). If $x(t) = s(t) + v(t)$ where $x(t)$ is a deterministic signal and $v(t)$ is a white noise, the simulation model can be found in [16], [24], [25]. There is a factor $\sqrt{\Delta t}$ to scale $v(t)$ such that the scaled $v(t)$ conforms with Winner increment, where $\Delta t$ is the sampling period in numerical simulation. While the $x(t)$ in Eq. (2), as will be seen later, is the watermark information embedded in the DCT coefficients and is not related to time. Therefore $\Delta t$ (the time interval of one sample applied to BS) in simulation can be chosen arbitrarily (but need to consider the $a$ and $b$ for a convergent output). $x(t)$ is the velocity added to the ball and the reasonable simulation model is [11]-[14], [17], [26]

$$p[n+1] = p[n] + \Delta t * (ap[n] - bp^3[n] + x[n]), \quad (3)$$

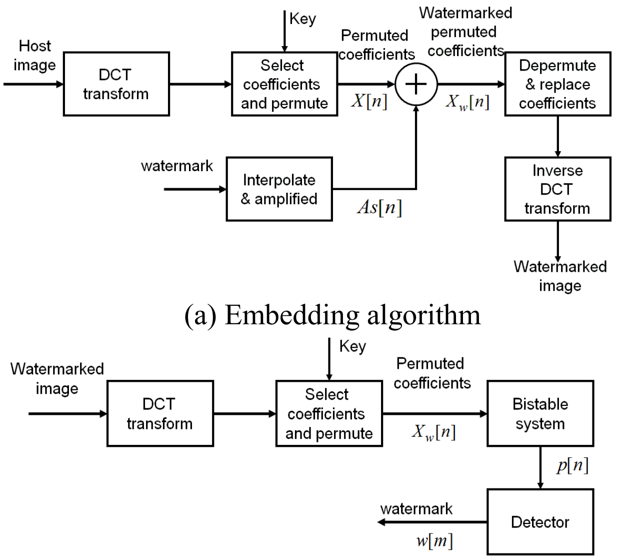
where $\sqrt{\Delta t}$ is sampling interval, and $p[n]$ is the position of the ball (i.e., the output of the system) at time index $n$ . If the parameters $a, b$ and $\Delta t$ are known, $p[n]$ can be calculated recursively using Eq. (3). A fourth-order Runge-Kutta (RK) method can be used to obtain a more precise result [13], [26].

## 2.2 BS-SR based watermark technique

As mentioned in Section 1, several watermarking techniques have been proposed using BS [11]-[14]. A watermarking technique includes two stages: embedding and extraction. In this section, we present a brief review of the watermark scheme in [13]. Note that the BS-SR detector is used only in the extraction.

### 2.2.1 Embedding algorithm

The schematic of the embedding algorithm is shown in Fig. 3 (a). The host image $I$ is divided into $8\times 8$ blocks. The DCT coefficients of $kth$ block are then calculated and zigzag scanned, and is denoted by $X^k, k=1,2,...,K$ . The middle $r$ DCT coefficients of $X^k$ are denoted by $X_u^k (U_1 \le u \le U_2)$ where $U_1$ is the starting index, $U_2$ is the ending index and $r = U_2 - U_1 + 1$. $X_u^k$ 's are then cascaded to generate a sequence $X$ with length $rK$ .




(a) Embedding algorithm

(b) Extraction algorithm

Fig. 3 Schematic of the SR watermarking scheme.

To make $X$ look like a white Gaussian noise (WGN), $X$ is permuted as follows. First, generate a random sequence $R$ with the same length of $X$ using a specific key. Second, generate $R'$ by sorting $R$ ascendant. Let $L$ contains the original index in $R$ of $R'$ elements. Finally, $X'$ is obtained via permuting $X$ by $L$ . For example, $X$=10,65,43,20, $R$=0.3,-0.1,-0.2,3, $R'$=-0.2,-0.1,0.3,3, $L$=3,2,1,4, $X' = 43,65,10,20$.

Assume $w[m], m \in [1, M]$ is a binary watermark sequence consisting of -1 and 1. Note that the binary watermark should be a sequence of 0 and 1 and 0 is converted to -1 for convenience. Every bit in $w[m]$ is repeated $S$ times and sequence $s[n]$ with the same length of $X'$ is generated. The watermark embedding algorithm is given by $X_w[n]=X[n]+As[n]$ , where $A$ is the watermark intensity factor, $X_w$ is the watermarked DCT coefficients. The original DCT coefficients are replaced by $X_w$ and then inverse DCT transform is performed to generate the watermarked image.

### 2.2.2 Extraction algorithm

The schematic of the watermark extraction is shown in Fig. 3 (b). The first two steps are identical to those of embedding scheme and obtain $X_w$ . We apply $X_w$ to the BS and obtain the output $p[n]$. Then $p[n]$ is partitioned into $M$ segments, each of which include $S$ samples corresponding to one bit of the watermark. The estimated position in the double well for one watermark bit is calculated as follows,

$$\mu[m] = \sum_{n=(m-1)*S+1}^{m*S} p[n] \ . \tag{4}$$

The watermark bit $w[m]$ is recovered by,

$$w[m] = \begin{cases} 1 & \mu[m] \ge 0 \\ -1 & \mu[m] < 0 \end{cases} \tag{5}$$

## 3 Proposed method for optimal parameter determination

The parameter set $(a, b, \Delta t)$ is crucial for the BS-SR detector to have a good performance. However the parameter set is difficult to determine, which limits the usage of the BS-SR detector in applications. In this section, we first analyze the influence of the parameters $a$ , $b$ and $\Delta t$ separately and reduce the determination of 3 parameters to only 1 parameter. We then show that the near-optimal parameters that make the BS have optimal SR effect are not sensitive to the pdf of the noise, only related to the variance of the noise and the amplitude of the signal. That leads to an empirical method to find the near-optimal BS parameters.

### 3.1 The influence of the parameters $a$, $b$ and $\Delta t$

First, it is worth clarifying the physically meaningful SR measure. If the optimal parameters are obtained when the overall detection performance, such as probability of detection, bit error rate (BER), is optimal, the BS may not work as a stochastic resonator [20]. Consider the function of the BS. The BS input $x[n]$ is composed of signal $s[n]$ and noise $v[n]$ and the BS output $p[n]$ can be considered as signal $s'[n]$ plus noise $v'[n]$ as well. The ideal (optimal) state of the SR effect is the signal $s'[n]$ should be like the signal $s[n]$ as much as possible. Therefore, the cross-correlation can be used as a SR measure, which is defined as follows,

$$C_m = E\left(\frac{\sum_{n=0}^{N-m-1} x[n]p[n+m]}{\sqrt{\sum_{n=0}^{N-m-1} x^2[n]}\sqrt{\sum_{n=0}^{N-m-1} p^2[n+m]}}\right) \quad (6)$$

where $m > 0$ is the time lag, $E(\cdot)$ denotes an average over the multiple experiments for a given $m$. We define $\tau$ as the system lag of the BS, which can be found by,

$$\tau = \underset{\tau=m}{\operatorname{argmax}} C_m . \quad (7)$$

With a given set of parameters $(a,b,\Delta t)$, the BS system has a delay $\tau$ and the cross-correlation is $C_\tau$.

Secondly, we show the influence of the parameters $a$, $b$ and $\Delta t$ below.

1) **The influence of $(a,b)$.** Parameters $a$ and $b$ are combined to decide the size of the double well as shown in Eq. (1). Assuming that $\Delta t$ is fixed, the signal $s[n]$, the noise $v[n]$ and parameters $(a,b)$ cooperate to have the near-optimal SR effect. In general, for a strong noise, the size of the double well shall be large. For a given $a$ (denoted as $a_1$), there shall be a suitable $b$ ($b_1$) that leads to a maximum cross-correlation. While for another $a$ ($a_2$), there shall be another suitable $b$ ($b_2$). If the BS with parameter sets $(a_1,b_1)$ and $(a_2,b_2)$ can have same (or almost same) cross-correlation, we can fix one parameter and only tune the other one. A justification is shown in Fig. 4. It is observed that for $(a,b) = (1,1000)$ and $(10,5000)$, the cross-correlations are same. We note that the cross-correlation between the input $x[n] = s[n] + v[n]$ and signal $s[n]$ is about 0.1 and the cross-correlation between the BS output and $s[n]$ is about 0.45 (as shown in Fig. 4), which means the signal $s[n]$ is enhanced in the view point of the SR measure.

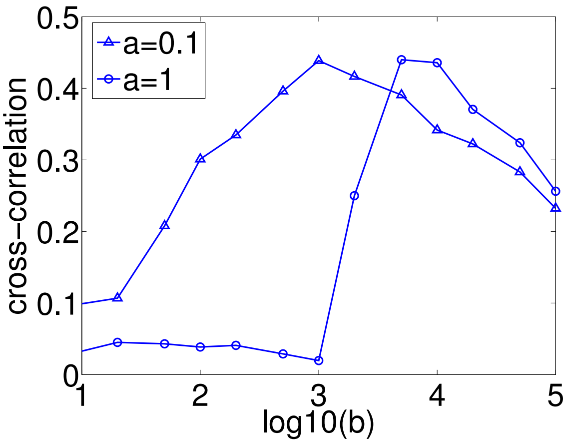



Fig. 4 Cross-correlation versus parameters $(a,b)$. In the simulation, signal is a 20 bit random binary PAM with $A = 0.1$. Every bit in the PAM signal is interpolated 50 times to generate a sequence $s[n]$ with length of 1000 samples. $s[n]$ is embedded in the 1000 i.i.d. WGN (zero-mean, unit variance) samples to generate the input to the BS. The BS parameters are with $\Delta t = 0.01$, $b$ is tuned to find the maximum cross-correlation for $a = 1$ and $a = 10$ respectively. Cross-correlation is calculated from 1000 simulations for a given $(a,b)$ using Eqs. (6) and (7).

2) **The influence of $\Delta t$.** $\Delta t$ is the time interval for one discrete time sample applied to the system. For a given $(a,b)$, if $\Delta t$ is too large, the system can always reach its stable position for a given sample; while if $\Delta t$ is too small, one sample has very little influence to the system output. It can be shown that the two extremes (too large or too small $\Delta t$) cause poor cross-correlation and $\Delta t$ shall match $(a,b)$ for a certain input. Fig. 5 shows that for a given $\Delta t$, there will be a suitable $(a,b)$ to have a maximum cross-correlation. When $\Delta t$ and $(a,b)$ are matched, for different $\Delta t$, the cross-correlations are quite close. In Fig. 5,

$(a,b,\Delta t)$ =(1,1000,0.01) and (1,10000,0.005) lead to the same cross-correlation.

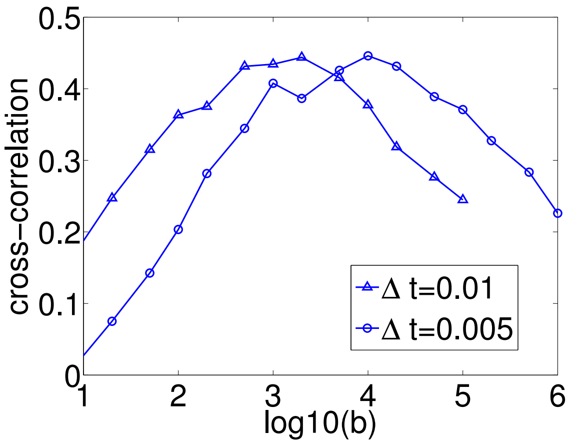



Fig. 5 Cross-correlation versus $\Delta t$. The setup of the experiment is the same as described in Fig. 4. The BS is with $a=1$ and $b$ is tuned for maximum cross-correlation for $\Delta t$ =0.01 and $\Delta t$ =0.005, respectively.

Thirdly, we propose a simplified tuning technique to determine the BS parameters. From the relationship between $(a,b)$ and the cross-correlation and the relationship between $\Delta t$ and the cross-correlation, we can conclude that $(a,b)$ and $\Delta t$ cooperate with each other to have the optimal SR effect (maximum cross-correlation). Because for different $a$, there shall be different $b$ for the maximum cross-correlation, and the maximum cross-correlations are very close for these $(a,b)$'s. Therefore, we can fix $a$ and only tune $b$ for the maximum cross-correlation. Similarly, different $\Delta t$ will have different $(a,b)$ and the maximum cross-correlation obtained from the different $\Delta t$ with the suitable $(a,b)$ are quite close. Hence, $\Delta t$ can be fixed a priori. Based the discussion above, the tuning technique in determining $(a,b,\Delta t)$ of the BS can be reduced from a 3-D tuning to 1-D tuning. First, we fix $a$, such as $a=1$. Second, determine a suitable $\Delta t$ according to the application. Suitable $\Delta t$ means that $b$ shall be reasonable (not very large and not very small) if this $\Delta t$ is applied to the BS. For example, $\Delta t = 0.01$, we find that $b=1000$ for maximum cross-correlation and $\Delta t$ is considered suitable. Finally, $b$ is tuned for the maximum cross-correlation. We only tune one parameter (TOP) in this tuning method, and thus TOP is used afterwards.

### 3.2 Near-optimal BS parameter determination

Even though we can use 1-D tuning to determine the BS parameters, the tuning technique is difficult to be put into applications because the pdf of the noise cannot be obtained accurately and we need the pdf of noise to find the BS parameters. Another important feature of the BS used as stochastic resonator is that the BS parameters are not sensitive to the pdf of the noise. Fig. 6 illustrates the SR measure versus the BS parameters and different noise pdfs. It is observed that BS parameters for maximum cross-correlation are same for the 4 pdfs with same variance $\sigma^2$, where the pdf of Gaussian mixture is defined as

$$p(x) = \frac{c}{\sigma\sqrt{2\pi}}\left[\alpha \exp\left(-\frac{c^2x^2}{2\sigma^2}\right) + \frac{1-\alpha}{\beta}\exp\left(-\frac{c^2x^2}{2\beta^2\sigma^2}\right)\right]$$

where $c = [\alpha + (1-\alpha)\beta^2]^{1/2}$, $0<\alpha<1, \beta>0$.

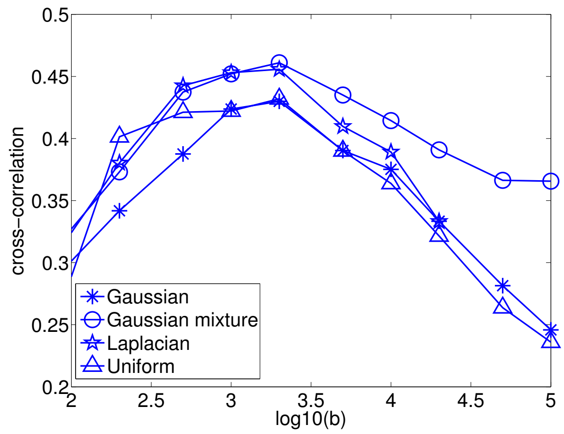



Fig. 6 Cross-correlation versus $b$ for different noise pdfs. The setup of the experiment is the same as described in Fig. 4. The BS is with $a=1$, $\Delta t$ =0.01 and $b$ is tuned to obtain maximum cross-correlation for different pdfs of the noise, which are Gaussian, Gaussian mixture, Laplacian and uniform. All the pdf are zero mean and variance $\sigma^2 = 1$. The Gaussian mixture is with $\alpha = 0.9$ and $\beta = 5$.

We can present a justification below. The suitable BS parameters mean that the physical size of the BS match the signal and the noise well. Consider that the evolvement of the ball in the double well. The ball is forced by $x[n]$ that is composed of the signal and the noise. From the velocity of the ball as shown in Eq. (2), we can think somehow that the power of the signal and the noise is crucial for the movement of the ball.

Assuming that the pdf of the noise is zero-mean and symmetric, the SR effect means that the position of the ball is in "resonance" with the signal. The most important factors in the best matching between the signal, the noise and the double well are the power of the signal, the power of the noise and the double well. The noise with large power will force the ball mount the central barrier frequently while a noise with small power will not. Therefore the characteristic of the evolvement of the ball is likely to be related to the power and not to the pdf of the noise.

This feature, the BS parameters are not sensitive to the pdf of the noise for optimal SR effect, is important to develop BS based applications. When the BS is used as a stochastic resonator, the near-optimal parameters will be identical for varying symmetric pdfs with same power. That provides an easy way to determine the BS parameters off-line because we can use any noise with zero-mean and symmetric pdf (the simplest one is WGN) to obtain the BS parameters by simulation. Combined with the TOP technique, we can construct a look up table (LUP) for the BS parameters with respect to $A$ (amplitude of watermark) and $\sigma^2$ (power of the noise). The procedure is shown below.

1) Generate the input by adding a random PAM (amplitude $A$) to a WGN (variance $\sigma^2$).
2) The BS is with fixed $(a, \Delta t)$ and $b$ is tuned to get maximum cross-correlation for the given input.
3) Go back to step 1 to generate the input with different $A$ and/or $\sigma^2$ and to find the appropriate $b$ in step 2.
4) The mappings between the parameters $b$ (with fixed $a$ and $\Delta t$) and $(A, \sigma^2)$ are saved in a LUT for later use.

Next we will show how to use the LUT to construct an adaptive BS-SR detector used in watermark extraction.

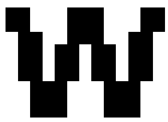
# 4 Proposed adaptive BS-SR detector

In Section 3, we proposed a method to determine the BS parameters. In this section, we consider the application of the method in digital watermarking and propose an adaptive BS-SR detector to extract the watermarks. We first analyze the extraction problem based on the embedding algorithm mentioned in Section 2.2 and an optimal ML detector is derived. We then show that in watermarking applications, designing an optimal detector is difficult because the noise pdf is not Gaussian and is difficult to estimate. Instead, a BS-SR detector could be employed as a suboptimal detector. Thirdly, we present the statistics of the input samples and show that the performance obtained by the BS with fixed parameters could be improved. Finally, an adaptive BS-SR detector is proposed.

## 4.1 The analysis of the extraction problem

In Section 2.2, we presented the approach to embed the watermark information. Let $A=3$, $U_1=9$, $U_2=44$, $M=256$ and $S=500$. The host Lena image and the binary watermark image are shown in Fig. 7.

(a) $512 \times 512$ Lena image



(b) $16 \times 16$ binary watermark image

Fig. 7 Lena image and embedded watermark image.

The watermarked DCT coefficients $X_W$ has 256 segments with every segment having 500 samples. The 500 samples in one segment can be looked at as adding 3 (if watermark is 1, $A=3$) or -3 (if watermark is 0) to 500 permuted DCT coefficients considered as i.i.d. (identical and independent distributed) WGN. Therefore, the extraction is a detection problem of two hypotheses, shown as follows.

$$\begin{cases} H_0 & x[n] = -A + v[n], n = 1,2,..., S \\ H_1 & x[n] = A + v[n], n = 1,2,..., S \end{cases}, \quad (8)$$

where $A$ is watermark intensity factor, $v[n]$ is the i.i.d. WGN (with zero-mean and variance $\sigma^2$). With the Bayesian paradigm, the probability of error is estimated as follows:

$$P_e = P(H_0 \mid H_1)P(H_1) + P(H_1 \mid H_0)P(H_0) \quad (9)$$

where $P(H_i \mid H_j)$ is the conditional probability that indicates the probability of deciding $H_i$ when $H_j$ is true. Our goal is to design a detector that minimizes $P_e$. In the watermark application, the prior probability, $P(H_0)$ and $P(H_1)$, is assumed identical, i.e., $P(H_0)=P(H_1)=1/2$ due to the blink watermark. Further assuming that the pdfs $p(x \mid H_0)$ and $p(x \mid H_1)$ are known, we can have a maximum likelihood (ML) detector that would decide $H_1$ if

$$\frac{p(x \mid H_1)}{p(x \mid H_0)} > 1, \quad (10)$$

which is the optimal detector. Provided that $p(x \mid H_1)$ and $p(x \mid H_0)$ are Gaussian, for the mean-shift Gaussian detection formulated in Eq. (8), it is well-known that we should decide $H_1$ if the test statistic $T(x)$ is:

$$T(x) = \sum_{n=1}^{S} x[n] > 0 \cdot \quad (11)$$

The pdf of $T(x)$ is also Gaussian. Therefore,

$$\begin{aligned} P_e &= \frac{1}{2}\left[P(T(x) > 0 \mid H_0) + P(T(x) < 0 \mid H_1)\right] \\ &= Q\left(\frac{A}{\sqrt{\sigma^2 / S}}\right) \end{aligned} \quad (12)$$

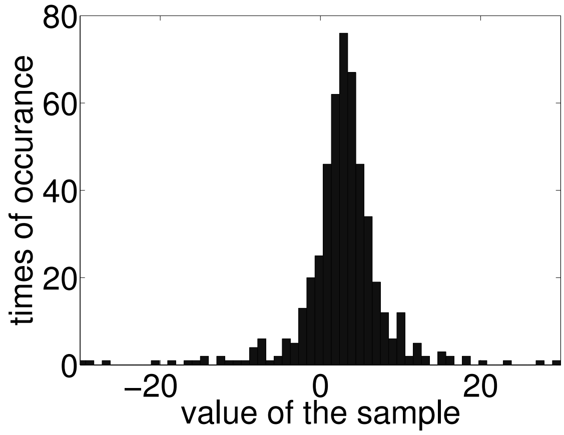
where $Q(\cdot)$ is the complementary cumulative distribution function (CDF) of the normal distribution. For $A = 3$ and $\sigma^2 = 100$, $P_e = 9.85 \times 10^{-12}$. In other words, the error occurs very rarely.

## 4.2 Optimal detection versus BS-SR detection

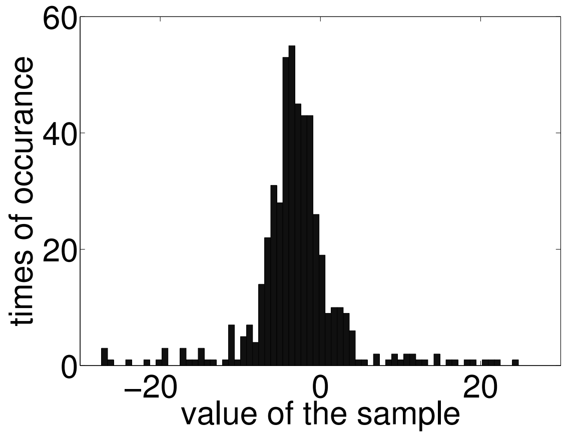
It is well known that the ML detector is an optimal detector in the error minimization sense when the noise is known. In particular, if the noise is Gaussian, the ML detector is linear as shown in Eq. (11). While for non-Gaussian noise, the ML detector is non-linear. Why do we employ the BS-SR detector to extract the watermarks? The reasons are listed as follows. To obtain the ML test statistics $T(x)$, we need to have the complete knowledge of the pdfs $p(x \mid H_0)$ and $p(x \mid H_1)$, which is not always available in practical applications. Also, the pdfs always vary with time and applications. In addition, for some non-Gaussian noise, $T(x)$ could be too complicated though we know $p(x \mid H_0)$ and $p(x \mid H_1)$ exactly. Therefore, we can rarely have the ideal conditions to design the optimal ML detector. The ML detector based on (11) is optimal if the noise is i.i.d. Gaussian, but will be suboptimal if the noise is not i.i.d. Gaussian. Alternatively, the BS-SR detector will provide a suboptimal performance. In the following, we first show that pdf of the samples in a segment of $X_W$ looks like Gaussian but changed significantly after the watermarked image is attacked by JPEG compression and Gaussian noise. We then justify that the BS-SR detector can have a better performance than Gaussian based ML detector in watermark extraction from JPEG compressed watermarked images.

Figure 8 shows the histograms of the samples of two segments (the 1st and 161th segments from watermarked Lena image) of $X_W$ corresponding to a binary watermark with values $A$ and $-A$. The histograms of the samples of the same two segments after the original watermarked image is attacked by JPEG compression and Gaussian noise are shown in Fig. 9 and Fig. 10, respectively. Note that the JPEG compression is performed by the Matlab function “imwrite(.)” with format “JPEG” and quality, such as 50 in Fig. 9. The Gaussian noise is added into the watermarked image by using the Matlab function “imnoise(.)” with the variance, such as 0.03 in Fig. 10. Although the pdfs of samples from the original watermarked image look like Gaussian (Fig. 8), but after JPEG compression, the pdfs do not look like Gaussian, (Fig. 9). That is because when the watermarked image is being compressed, many DCT coefficients are truncated to zero, which means that both the watermark information embedded and the original DCT coefficients are degraded. Therefore, we observe that many samples are zeros and the other samples are quite scattered. If we add

Gaussian noise to the watermarked image, the mean does not change a lot but the variance is about 30 times the variance of the samples from the original watermarked image. The pdfs are shown in Fig. 10.

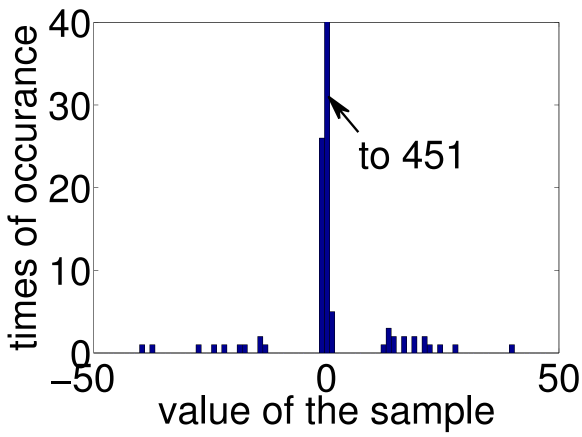



**(a)** watermark bit = 1

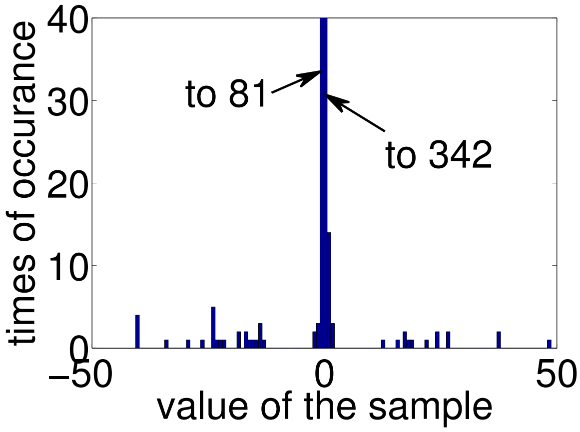



**(b)** watermark bit = -1

Fig. 8 The histograms of two segments in $X_W$ of the original watermarked image.

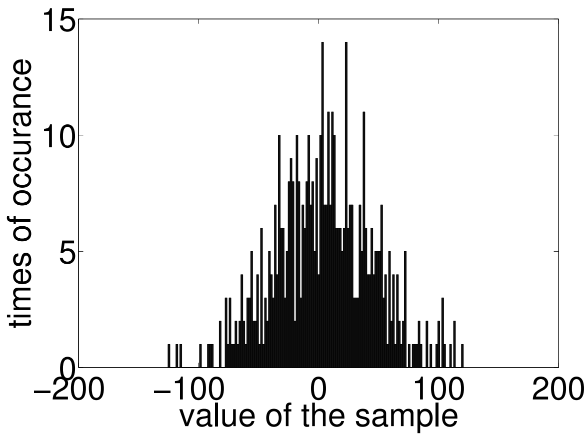



(a) watermark bit = 1

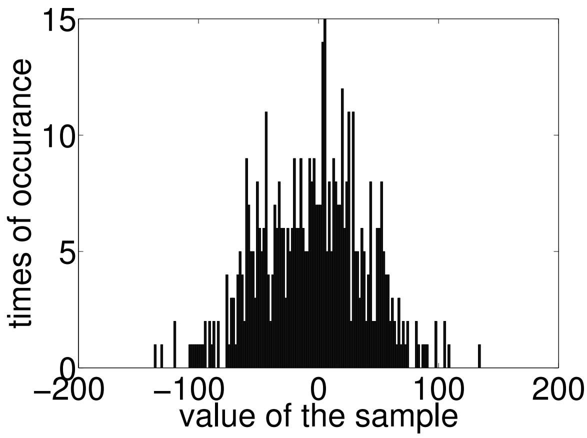



(b) watermark bit = -1

Fig. 9 The histograms of two segments in $X_W$ of the JPEG compressed watermarked image with quality of 50..




(a) watermark bit = 1

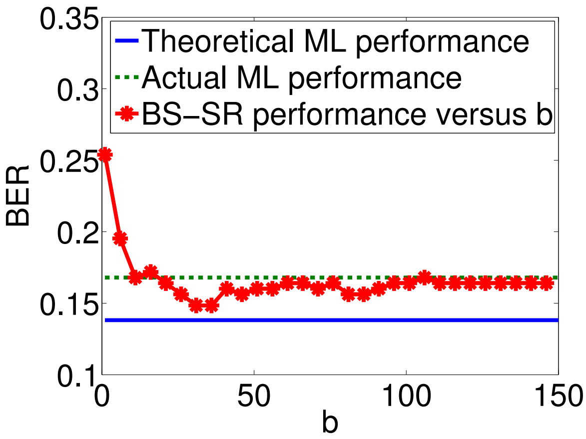



(b) watermark bit = -1

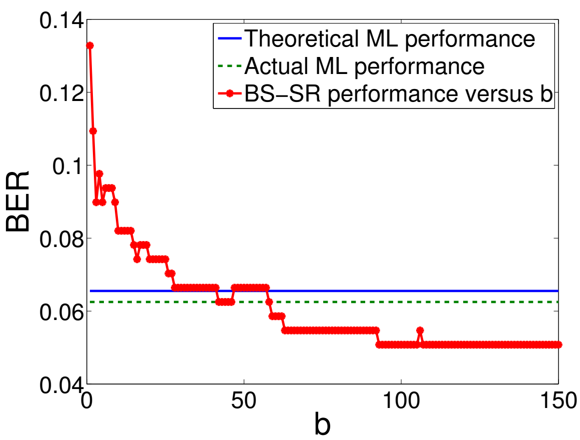
Fig. 10 The histograms of two segments in $X_W$ of the watermarked image added Gaussian noise with variance of 0.03.

We show that the samples in one segment of $X_W$ corresponding to one watermark bit may not be Gaussian and it is usually difficult to estimate the exact pdf of these samples. Therefore, the optimal ML detector is difficult to derive and the ML detector based on Eq. (12) (still called ML detector afterwards) will be a suboptimal detector. A BS-SR detector with appropriate parameters may improve the performance. However the choice of the parameters is critical to obtain a good performance.

The performance of the BS-SR detector and the ML detector are compared in Fig. 11 with two examples. The solid line is the theoretical ML performance calculated by Eq. (13). The dotted line is the real performance of the ML detector. It is obvious that the BS-SR detector with suitable parameters can have a better performance. We have shown that in practice, it is not feasible to obtain good BS parameters by tuning. The method presented in Section 3 will be used to build a BS-SR detector with adaptive BS parameters.




(a) Quality 30



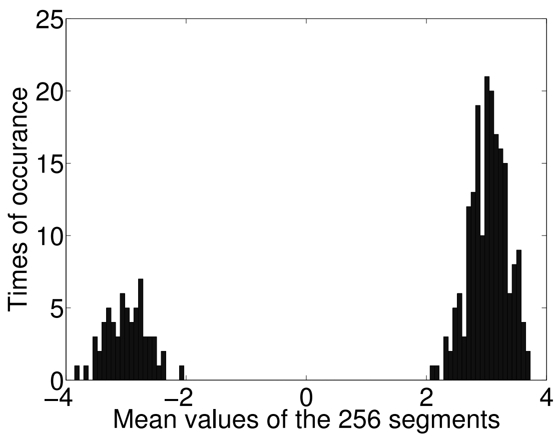

(b) Quality 50

Fig. 11 The BERs of the extracted watermark from watermarked Lena image (subject to JPEG compression with quality 50 and 30) against parameter $b$ of the bistable system with $a = 1$ and $\Delta t$ =0.01.

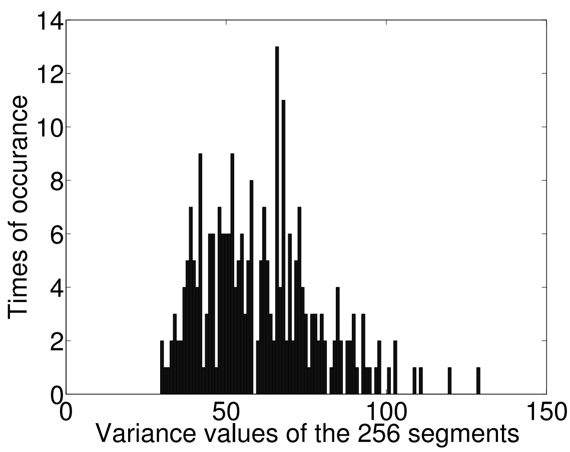
### 4.3 Nature of watermarked DCT coefficients $X_W$

We present the statistical features of the $X_W$, from which the watermark is extracted. And the reasons of using the adaptive BS-SR detector are shown as follows.

We embed a watermark image into Lena image and obtain $X_W$ as described in Section 4.1. Totally there are 256 segments in $X_W$, every segment includes 500 samples corresponding to one watermark bit. Now we present the statistical features of the 256 segments in Fig. 12, Fig. 13 and Fig. 14.

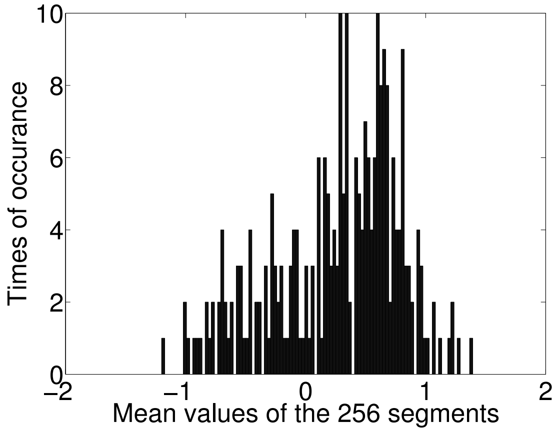



(a) Mean

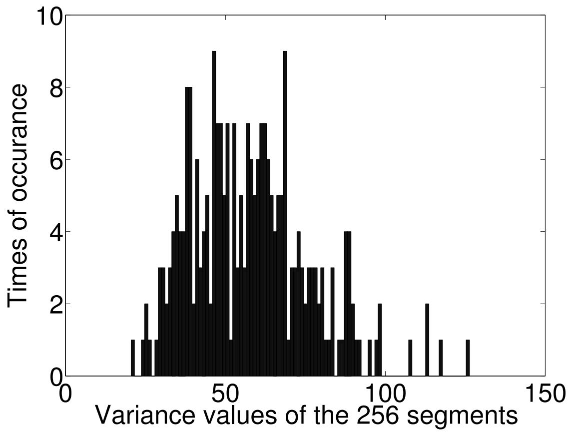



(b) Variance

Fig. 12 The histograms of the mean and variance of the 256 segments in $X_W$ of the original watermarked Lena image.

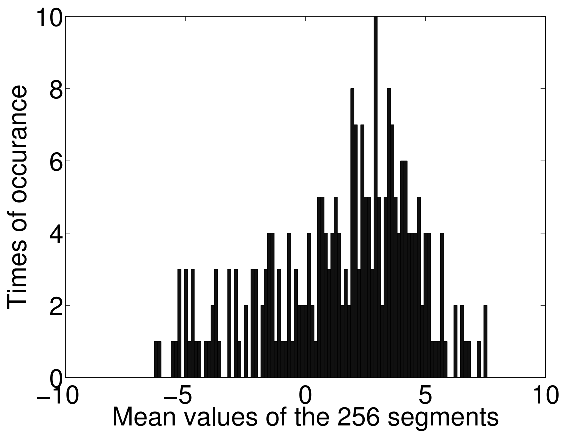


(a) Mean

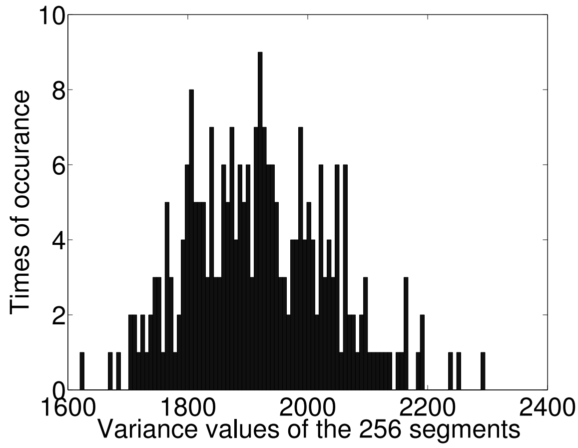


(b) Variance

Fig. 12 The histograms of the mean and variance of the 256 segments in $X_W$ of the watermarked Lena image subject to JPEG compression with quality of 50.



(a) Mean



(a) Variance

Fig. 14 The histograms of the mean and variance of the 256 segments in $X_W$ of the watermarked Lena image subject to the Gaussian noise attack with variance of 0.03.

Figure 12 shows the means are near 3 and -3 and the variances vary from 30 to 140. It is easy to divide the watermark embedded into two categories. The means near -3 represent the watermark bit -1 and the means near 3 represent the watermark bit 1. There is a big gap between these two categories and therefore it is easy to make a decision. However, when the watermarked image is attacked by JPEG compression or by adding Gaussian noise, we show that a lot of the watermark information is lost and the two categories become merged. For JPEG compression, the watermark information is filtered heavily, showing the means are very near 0, but the variances change little. For Gaussian noise attack, the watermark information is not lost, but the noise is strong and masks the watermark information. We can consider that the mean of the samples in a segment corresponding to one bit of the watermark denotes the strength of the watermark $A$ varying from very weak (such as 0.1) to very strong (such as 3). The variance represents the noise strength $\sigma^2$, and also varies from very weak (30) to very strong (2300).

From the discussion in Section 3, we know a suboptimal parameter set $(a, b, \Delta t)$ depending on ($A$, $\sigma^2$), and there should be different $(a, b, \Delta t)$ for the BS to have a good SR effect. It would be difficult to obtain a good performance by providing only one set $(a, b, \Delta t)$ for significantly varied ($A$, $\sigma^2$). Therefore we propose an adaptive BS-SR detector to achieve better performance.

### 4.4 The proposed adaptive BS-SR detector

A suitable $(a, b, \Delta t)$ could be decided for a given $A$ and $\sigma^2$. Instead of using one $(a, b, \Delta t)$ for the various images [12], [13], we employ adaptive parameters for various images or even for various segments of an image. We describe the adaptive BS-SR detector and use it in the watermark extraction.

For watermark scheme, the embedding algorithm is the same as that shown in Fig. 3 (a). The decoding algorithm is similar to that shown in Fig. 3 (b), except the bistable system module. The bistable system module in Fig. 3 (b) is changed to adaptive bistable system and its schematic is shown in Fig. 15. It is observed that there are 4 sub-modules in the system. These sub-modules are explained below.

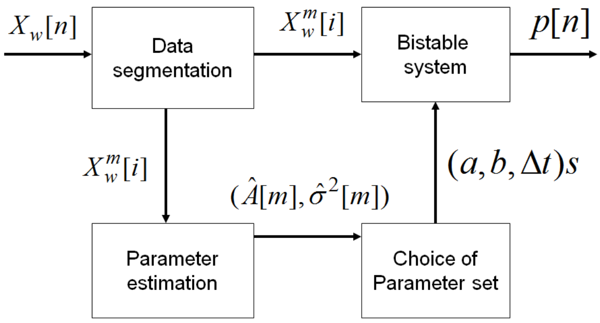



Fig. 15 Schematic of adaptive bistable system.

1) **Data segmentation:** $X_w[n]$ is partitioned into $M$ segments, which are denoted as $X_w^m[i], m \in [1, M], i \in [1, S]$ corresponding to $M$ bits of the watermark, where $X_w^m[i] = X_w[(m-1) * S + i]$.

2) **Parameter estimation:** We model $X_w^m[i]$ as a watermark embedded in i.i.d. WGN and estimate the mean $\hat{A}[m]$ and variance $\hat{\sigma}^2[m]$. Note that $\hat{A}[m]$ and $\hat{\sigma}^2[m]$ are the maximum likelihood estimation (MLE) of the embedding amplitude $A$ and the noise variance $\sigma^2$ in the $mth$ segment, which are shown as

$$\hat{A}[m] = \frac{1}{S}\left|\sum_{i=1}^{S} X_w^m[i]\right| \tag{13}$$

$$\hat{\sigma}^2[m] = \frac{1}{S}\sum_{j=1}^{S}\left(X_w^m[j] - \frac{1}{s}\sum_{i=1}^{S} X_w^m[i]\right)^2 \tag{14}$$

3) **Choice of the parameter set:** We choose an appropriate $b$ from the LUT according to the estimated $(\hat{A}, \hat{\sigma}^2)$. With the fixed $a$ and $\Delta t$, $(a, b, \Delta t)$ are applied to the BS.

4) **Bistable system:** The input sequence is processed by the BS with adaptive parameters and generate the output $p[n]$, which is used to extract watermark information according to Eqs. (4) and (5).

We note that the proposed BS-SR detector uses the parameters adaptive to the every segment $X_w^m[i]$, which means for every segments, the $\hat{A}[m]$ and $\hat{\sigma}^2[m]$ are estimated and the suitable parameters are chosen from the LUT. In watermark extraction, we find that the BS parameters are not very sensitive to the extraction performance, and therefore, we can only estimate one $(\hat{A}, \hat{\sigma}^2)$ for one image (not for every segment) to reduce the implemental complexity. The MLE of the $(\hat{A}, \hat{\sigma}^2)$ of one image is shown below:

$$\hat{A} = \frac{1}{M}\left|\sum_{m=1}^{M} \hat{A}[m]\right| \tag{15}$$

$$\hat{\sigma}^2 = \frac{1}{M}\sum_{m=1}^{M} \hat{\sigma}^2[m] \tag{16}$$

where $\hat{A}[m]$ and $\hat{\sigma}^2[m]$ are the estimated watermark amplitude and variance of the $mth$ segment as calculated in Eqs. (13) and (14).

We name BS-SR-I for the BS-SR detector that only chooses the BS parameters once for one image (using Eqs. (15) and (16)). We name BS-SR-II for those that choose the BS parameters for every segment. Compared to the existing SR scheme, our scheme chooses parameters adaptively based on the $(\hat{A}, \hat{\sigma}^2)$ estimated from varying images or even varying segments of one image. The performance should be better than that obtained from that with fixed parameters. In addition, the determination of the parameters is more practical than the techniques by tuning parameters.

## 5 Experimental results

In this section, experimental results are given to illustrate the transparency, robustness against JPEG compression and Gaussian noise. We omit other experiments for simplicity and the results provided are enough to arrive at a conclusion of the BS-SR detection performance. The images, Lena, Peppers, Goldhill and Baboon are chosen as the host images. We embed watermark shown in Fig. 7 (b) into host images by the embedding algorithm presented in Section 2.2 with $A = 3$, $U_1 = 9$, $U_2 = 44$, $M = 256$ and $S = 500$. We present the performance results from Gaussian based ML detector, Wu's method [13], Sun's method [11], BS-SR I detector and BS-SR II detector. The PSNRs of the watermarked images are all 41.07 dB.

First, we show the performance after the watermarked images are attacked by JPEG compression. The BERs are shown in Table 1 with respect to the five techniques and the compression qualities (from 80 to 30). Note the BERs are all 0 when the compression quality is 100 and 90 for the five techniques. From Table 1, it is observed that the performance of ML detector is always better than Wu's method, which means that Gaussian based ML detector works well as a suboptimal detector. Sun's method can provide a better performance than the ML detector. The performance of the proposed BS-SR I and BS-SR II detectors are shown in the two rightmost columns. It is observed that the performance of the BS-SR detectors are generally better than the ML detector and Sun's method and the adaptive BS-SR detector is suitable for all kinds of images. Comparing BS-SR I and BS-SR II, we found that the latter provides only a little improvement but not significantly.

Secondly, we add Gaussian noise to the watermarked images. The results are shown in Table 2. The ML detector has a very good performance. For the 4 images in our experiments the performance of the ML detector is always better than that of Wu's method. Compared with the ML detector, the BS-SR detector is as good as the ML detector in term of BER.

We show that the BS-SR detector has an improved performance compared to the ML detector when the watermarked image is attacked by JPEG compression. However, the BS-SR detector may not have a better performance than the ML detector when the image is attacked by Gaussian noise. The reasons are depicted as follows. First, if the noise is i.i.d. WGN, the ML detector is the optimal detector. When the image is compressed by JPEG, many coefficients are converted to 0 that leads to the loss of the signal (watermark) as well as the noise. The coefficients in one segment may not be a WGN, which is shown in Fig. 13 (a) and (b). The Gaussian based ML detector is not optimal detector and hence, the BS-SR detector employed as suboptimal detector provides an improved performance. Secondly, when adding Gaussian noise to an image, the signal information will not loss but the noise will become very strong. The pdf of the coefficients is still considered as Gaussian, see Fig. 14 (a) and (b). We can still consider that the watermark is embedded in WGN, so the ML detector tends to be the optimal detector. That is why the BS-SR cannot have an improved performance. However, the BS-SR detector can be fairly compared to the ML detector.

## 6 Conclusions and discussions

The motivation for this research is to show how we can use the BS-SR detector for the binary PAM signal embedded in non-Gaussian noise. The signal can be extended for periodic signals or any other ones that can be converted to two states, and hence the proposed BS-SR could be used for other applications, such as sinusoidal signal detection in non-Gaussian noise. To simplify the choice of the parameters of the BS-SR detector, based on the effect of the BS parameters, a TOP technique was proposed to reduce the 3-D tuning to 1-D tuning. Furthermore, it was shown that the SR effect in the BS is not sensitive to the pdf of the noise, which leads to an empirical method to determine the BS parameters by an off-line simulation. In addition, an adaptive SR detector is constructed via choosing suitable parameters for the everchanging input. For the BS based watermarking scheme, we can provide a method to determine the BS parameters by an off-line simulation and the suitable BS parameters can be chosen via estimating the amplitude of the signal and the amplitude of the noise when extracting the watermarks. We show that the adaptive BS-SR detector can achieve a better performance than a linear detector. When pdf is difficult to obtain and the optimal detector is unavailable, the adaptive BS-SR detector provides a suboptimal choice.

It is worth pointing out that our techniques can construct a BS-SR detector and have a stable performance, but the performance is not optimal. It was shown that the pdf of the DCT coefficients are non-Gaussian noise with heavy pdf tails [21]. For this kind of non-Gaussian noise, a good detector typically includes a nonlinear limiter to reduce the noise spikes [22]. The BS can provide a suitable

**Table 1**. The BERs (in %) of the extracted watermark subject to JPEG compression with different qualities. The BERs in bold font are the minimum values for a specific case. The parameters used in Wu's method are $(a,b)=(4000,3\times10^{11})$ and $\Delta t=10^{-6}$. The parameters used in Sun's method are $(a,b)=(500,3\times10^{10})$ and $\Delta t=10^{-5}$. For BS-SR I, the parameters used are $\Delta t=0.01$, $a=1$ and $b=[40,40,45,45,50,60]$ for Lena corresponding to the six compression qualities, $b=[30,30,40,40,30,45]$ for Peppers, $b=[15,15,15,15,30,30]$ for Goldhill and $b=[3,3,3,3,3,3]$ for Baboon. For BS-SR II, different $b$ is chosen for every segment with fixed $\Delta t=0.01$ and $a=1$.

| Image | Quality | Methods | | | | |
|---|---|---|---|---|---|---|
| | | ML method | Wu's method | Sun's method | BS-SR I | BS-SR II |
| Lena | 80 | **0** | 0.39 | **0** | **0** | **0** |
| | 70 | **0** | 0.43 | **0** | **0** | **0** |
| | 60 | **1.56** | 9.38 | **1.56** | 1.95 | 1.95 |
| | 50 | **6.25** | 14.45 | 6.64 | **6.25** | 6.64 |
| | 40 | 10.55 | 18.75 | 9.77 | 9.77 | **8.98** |
| | 30 | 16.8 | 26.56 | **14.45** | 16.41 | 16.41 |
| peppers | 80 | **0** | 1.56 | **0** | **0** | **0** |
| | 70 | 2.73 | 4.69 | 1.56 | 1.56 | **1.17** |
| | 60 | 7.03 | 10.16 | 2.34 | **1.56** | 1.95 |
| | 50 | 12.5 | 18.36 | **8.98** | 9.38 | **8.98** |
| | 40 | 18.75 | 25.39 | 16.41 | 16.41 | **15.23** |
| | 30 | 26.56 | 35.55 | 25 | **24.61** | 25 |
| Goldhill | 80 | **0** | 0.39 | **0** | 0.39 | 0.39 |
| | 70 | 0.39 | 1.56 | **0** | **0** | **0** |
| | 60 | 1.17 | 3.52 | **0.78** | **0.78** | **0.78** |
| | 50 | **1.95** | 5.47 | **1.95** | 2.34 | 1.95 |
| | 40 | 7.42 | 7.42 | **4.69** | **4.69** | **4.69** |
| | 30 | 13.67 | 14.06 | 10.16 | 10.55 | **9.77** |
| Baboon | 80 | 1.17 | 1.56 | 1.17 | 1.17 | **0.78** |
| | 70 | 1.95 | 1.95 | **1.56** | 1.95 | **1.56** |
| | 60 | **2.73** | 3.52 | **2.73** | 3.13 | **2.73** |
| | 50 | 4.3 | 6.25 | 4.3 | **3.91** | **3.91** |
| | 40 | 5.86 | 7.81 | 5.47 | **5.08** | **5.08** |
| | 30 | 10.94 | 11.33 | 11.72 | 10.16 | **9.77** |

**Table 2.** The BERs (in %) of the extracted watermark subject to the attacks of Gaussian noise with different variances. The parameters used in Wu's method and in Sun's method are identical to the parameters shown in Table 1. For BS-SR I, the parameters used are $\Delta t = 0.01$ , $a = 1$ and $b = [40, 20, 15, 10, 5]$ for Lena corresponding to the five variance of the additive WGN, $b = [40, 40, 20, 15, 10]$ for Peppers, $b = [40, 40, 15, 10, 10]$ for Goldhill and $b = [20, 15, 10, 5, 5]$ for Baboon. For BS-SR II, different $b$ is chosen for every segment with fixed $\Delta t = 0.01$ and $a = 1$.

| Image | Variance added | Methods | | | | |
|---|---|---|---|---|---|---|
| | | ML method | Wu's method | Sun's method | BS-SR I | BS-SR II |
| Lena | 0.01 | **0** | 0.39 | **0** | **0** | **0** |
| | 0.015 | **0** | 0.43 | **0** | **0** | **0** |
| | 0.02 | **1.56** | 9.38 | **1.56** | 1.95 | 1.95 |
| | 0.025 | **6.25** | 14.45 | 6.64 | **6.25** | 6.64 |
| | 0.03 | 10.55 | 18.75 | 9.77 | 9.77 | **8.98** |
| Peppers | 0.01 | **0.78** | 1.95 | 1.17 | 1.17 | 1.56 |
| | 0.015 | **2.34** | 3.91 | **2.34** | **2.34** | 3.13 |
| | 0.02 | **3.91** | 7.03 | 5.08 | 4.3 | 5.08 |
| | 0.025 | 8.2 | 8.2 | **6.25** | 7.42 | 7.03 |
| | 0.03 | **7.81** | 9.77 | 8.59 | 8.98 | 9.38 |
| Goldhill | 0.01 | **1.56** | **1.56** | **1.56** | **1.56** | **1.56** |
| | 0.015 | **2.34** | 3.52 | **2.34** | **2.34** | 2.73 |
| | 0.02 | 4.69 | **4.3** | 5.47 | 5.47 | 5.47 |
| | 0.025 | **5.47** | **5.47** | 7.81 | 6.64 | 6.64 |
| | 0.03 | **8.59** | 11.33 | 9.77 | 10.55 | 10.94 |
| Lena | 0.01 | **3.52** | **3.52** | **3.52** | **3.52** | 3.91 |
| | 0.015 | **5.08** | 6.25 | 6.64 | 6.25 | 6.25 |
| | 0.02 | **5.86** | 7.03 | 7.42 | 7.03 | 6.64 |
| | 0.025 | **5.86** | 7.42 | 6.25 | **5.86** | 7.03 |
| | 0.03 | 10.55 | 10.94 | **9.77** | 10.55 | 10.94 |

nonlinearity to be employed as this nonlinear limiter. The BS used as a limiter followed by a linear detector will be a locally optimal (LO) detector [23]. If we tune the BS parameters for an overall optimal BER, the BS is more likely to be a limiter and that is why it can have an improved performance compared to the ML detector. However, in this case, the BS is not employed as a stochastic resonator and it is difficult to put this technique into real applications. The method proposed in this paper determines the optimal parameters for the BS to be a stochastic resonator. We consider the BS separately and the BS parameters are chosen based on maximum SR measure (cross-correlation measure), which are not the parameters for minimum BER. Therefore, the optimal parameters for BS to be stochastic resonator is not the optimal BS parameters for the optimal BS-SR detector.

The second issue is why we cannot achieve the optimal performance even through the optimal parameters that lead to maximum cross-correlation are applied to the BS. That will also answer the question: why we cannot obtain a significant improvement compared to the ML detector and the other BS based detector. The reason is that the linear detector followed the BS used as stochastic resonator (see Fig. 1) is not optimal. Considering the output of the BS, we can think that it is composed of signal and noise. To design an optimal detector in the view point of the BS output, the pdf of the noise at the BS output side needs to be estimated. That is difficult because the BS output is Markov chain but not homogenous. We simply use linear detector followed the BS. If the inner detector is linear, for the non-Gaussian noise with heavy pdf tails, the BS shall be a suitable nonlinear limiter to construct a LO detector. The parameters for the BS to have an optimal SR effect are not those that make the BS to have a suitable nonlinearity. That is why the optimal parameters for the BS to be a stochastic resonator are not the parameters for the BS-SR detector to have an optimal BER. In a word, the linear inner detector for the BS used as a nonlinear limiter is optimal but is not optimal for the BS used as stochastic resonator. In this paper, we proposed several techniques to design the BS as a stochastic resonator, but did not consider the design of the inner detector followed the BS. The optimal design of the inner detector is an open question and the further improvement of the BS-SR detector depends on the progress of this question.

*References:*


[1] R. Benzi, A. Sutera, and A. Vulpiani, The mechanism of stochastic resonance, *Journal of Physics A: Mathematical and General*, vol.14, No.11, 1981, pp. L453-L457.

[2] F. Moss, L. Ward, and W. Sannita, Stochastic resonance and sensory information processing: a tutorial and review of application, Clinical Neurophysiology, Vol.115, 2004, pp. 267-281.

[3] F. Duan and B. Xu, Parameter-induced stochastic resonance and baseband binary pam signals transmission over an awgn channel, *International Journal of Bifurcation and Chaos,* Vol.13, No.2, 2003, pp. 411-425.

[4] B. Xu, F. Duan, R. Bao, and J. Li, Stochastic resonance with tuning system parameters: the application of bistable systems in signal processing, *Chaos, Solitons and Fractals*, Vol.13, No.4, 2002, pp. 633-644.

[5] B. Xu, F. Duan, and Chapeau-Blondeau, Comparison of aperiodic stochastic resonance in a bistable system realized by adding noiseand by tuning system parameters, *Phys. Rev. E*, Vol.69, No.6, 2004, pp. 061110.

[6] A. Saha and G. Anand, Design of detectors based on stochastic resonance, *Signal Process*, Vol.83, No.6, 2003, pp. 1193-1212.

[7] H. Chen, P. Varshney, S. Kay, and J. Michels, Theory of the stochastic resonance effect in signal detection:part I-fixed detectors, *IEEE Transactions on signal processing*, Vol.55, No.7, 2007, pp. 3171-3184.

[8] G. Guo, M. Mandal, and A. Saha, Improved method to detect weak sinusoidal from ocean acoustic noise based on stochastic resonance, *Proceedings of the 10th IASTED international conference on Signal and Image Processing*, 2008.

[9] H. Chen and P. Varshney, Theory of the stochastic resonance effect in signal detection: part II-variable detectors, IEEE Transactions on signal processing, Vol.56, No.10, 2008, pp. 5031-5041.

[10] R. Peng, H. Chen, and P. Varshney, Stochastic resonance: an approach for enhanced medical image processing, *2007 IEEE/NIH science systems and applications workshop*, 2007.

[11] S. Sun and P. Qiu, Algorithm of digital watermarking based on parameter-induced stochastic resonance, Journal on Communications, Vol.26, No.12, 2005, pp. 48-55.

[12] S. Sun and S. Kwong, Digital watermarking based on stochastic resonance signal processor, *PCM 2007,LNCS4810*, 2007, pp. 367-376.

[13] G. Wu and Z. Qiu, A novel watermarking scheme based on stochastic resonance, proceedings of the 8th International Conference on Signal Processing, 2006, pp. 16-20.
[14] F. Duan and D. Abbott, The application of nonlinear bistable detectors to DCT-domain watermarking schemes, *Procedings Of SPIE*, Vol.6802, No.15, 2008.
[15] D. Rousseau, J. Rojas, F. Duan, and F. Chapeau-Blondeau, Evaluation of a nonlinear bistable filter for binary signal detection, I*nternational Journal of Bifurcation and Chaos,* Vol.15, No.2, 2005, pp. 667—679.
[16] M. E. Inchiosa and A. R. Bulsara, Signal detection statistics of stochastic resonators, *Phys. Rev. E*, Vol.53, No.3, 1996, pp. R2021--R2024.
[17] V. Galdi, V. Pierro, and I. M. Pinto, Evaluation of stochastic-resonance-based detectors of weak harmonic signals in additive white gaussian noise, *Phys. Rev. E*, Vol.57, No.6, 1998, pp. 6470—6479.
[18] S. Zozor and P.-O. Amblard, On the use of stochastic resonance in sine detection, *Signal Processing,* Vol.82, No.3, 2002, pp. 353-367.
[19] F. Duan, D. Rousseau, and F. Chapeau-Blondeau, Residual aperiodic stochastic resonance in a bistable dynamic system transmitting a suprathreshold binary signal, *Phys. Rev. E*, Vol.69, No.1, 2004, pp. 011109.
[20] G. Guo and M. Mandal, Design of stochastic-resonator-based detector using bistable system, *IEEE Conference on SPCOM*, 2010.
[21] A. Briassouli and M. Strintzis, Locally optimum nonlinearities for dct watermark detection, IEEE Transactions on Image Processing, Vol.13, No.12, 2004, pp. 1604-1617.
[22] S. Kay, *Fundamentals of statistical signal processing, Volume II: Detection theory*, Prentice Hall PTR, 2008.
[23] G. Guo and M. Mandal, Design of bistable system based nonlinear detector, *IEEE Conference on MWSCAS,* 2010.
[24] T. C. Gard., *Introduction to stochastic differential equations,* M. Dekker, New York, 1988.
[25] S. Mitaim and B. Kosko, Adaptive stochastic resonance, *Proceedings of the IEEE*, Vol.86, No.11, 1998, pp. 2152-2183.
[26] A. S. Asdi and A. H. Tewfik, Detection of weak signals using adaptive stochastic resonance, *Proceedings of the ICASSP*, 1995, pp. 1332-1335.