Physics Letters A 379 (2015) 2277–2283

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

Decoding suprathreshold stochastic resonance with optimal weights

Liyan Xu [a,*], Tony Vladusich [b], Fabing Duan [a], Lachlan J. Gunn [c], Derek Abbott [c], Mark D. McDonnell [b,c]

[a] Institute of Complexity Science, Qingdao University, Qingdao 266071, PR China
[b] Computational and Theoretical Neuroscience Laboratory, Institute for Telecommunications Research, School of Information Technology and Mathematical Sciences, University of South Australia, SA 5095, Australia
[c] Centre for Biomedical Engineering (CBME) and School of Electrical & Electronic Engineering, The University of Adelaide, Adelaide, SA 5005, Australia

ARTICLE INFO

Article history:
Received 28 December 2014
Received in revised form 14 May 2015
Accepted 14 May 2015
Available online 19 May 2015
Communicated by C.R. Doering

Keywords:
Suprathreshold stochastic resonance
Binary quantizer
Mean square error distortion
Stochastic quantization
Least squares regression
Information transmission

ABSTRACT

We investigate an array of stochastic quantizers for converting an analog input signal into a discrete output in the context of suprathreshold stochastic resonance. A new optimal weighted decoding is considered for different threshold level distributions. We show that for particular noise levels and choices of the threshold levels optimally weighting the quantizer responses provides a reduced mean square error in comparison with the original unweighted array. However, there are also many parameter regions where the original array provides near optimal performance, and when this occurs, it offers a much simpler approach than optimally weighting each quantizer's response.

1. Introduction

The term stochastic resonance (SR) [1–5] is used to describe phenomena where improvement of transmission or processing of a signal in a nonlinear system is achieved by tuning the noise intensity. Since its origins thirty years ago in the field of geophysical dynamics [1], SR has received considerable attention in a growing variety of systems with various types of signals and performance measures [6–21]. Most SR studies carried out today occur in threshold-based or potential barrier systems where a signal is by itself too weak to overcome a threshold or a potential barrier [6–21], but the presence of noise allows the signal to cross the threshold eliciting a more effective system response. Therefore, subthreshold input signals in threshold-based systems were originally assumed to be a necessary condition for the occurrence of SR.

Interestingly, a form of SR was reported by Stocks [22–24], under the name of suprathreshold SR (SSR), since it operates with signals of arbitrary magnitude, not restricted to weak or subthreshold signals. Notably, SSR is an important extension of SR with potential applications in a range of areas including neural systems. For example, SSR has been considered in ensembles of sensory neurons [25], signal quantizers [41], cochlear implant devices [26] and nonlinear detectors [28]. Moreover, artificial sensors, digital beamforming, biological neurons, cochlear implants and multiaccess communication systems can all be unified under the concept of stochastic pooling networks that manifest the noise-enhanced processing property [32–34]. Due to the variety of scenarios where SSR is observed, a number of performance measures have been considered, for instance, mutual information [22,23,27,29–31], mean square error (MSE) distortion [35,41,43], input–output cross-correlation [35,38], Fisher information [36,39,43] and signal-to-noise ratio [37].

The model studied in [22] that exhibits SSR is effectively a stochastic quantizer, since it converts an analog input signal into a digital output signal with threshold values randomized by noise [40–43]. McDonnell et al. have analyzed SSR in terms of lossy source coding and quantization theory, and examined the optimality of the quantization by using MSE distortion [40–43]. It was shown that the case of all identical threshold values is optimal for sufficiently large input noise, and a bifurcation pattern appears in the optimal threshold distribution with decreasing noise intensity, whether maximizing the mutual information or minimizing the MSE distortion [40–43].

In this paper, we investigate the decoding scheme of a quantized signal in the generic SSR model [22]. We propose a new

* Corresponding author.
E-mail address: xuliyan@qdu.edu.cn (L. Xu).

http://dx.doi.org/10.1016/j.physleta.2015.05.032

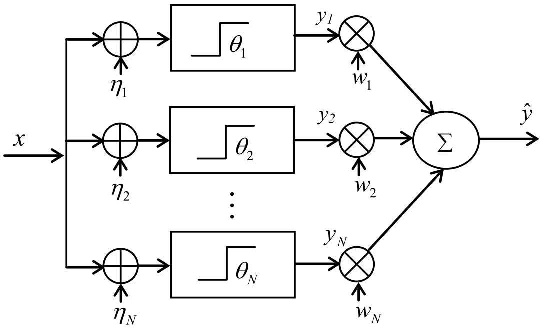

Fig. 1. Weighted summing array of N noisy comparators. It consists of N identical comparators (i.e. single bit quantizers), each operating on a common signal x subject to independent additive noise η_i. The output of each individual comparator, y_i, is multiplied by the weighted coefficient w_i, resulting in the weighted output $w_i y_i$. The overall output, $\hat{y}$, is the sum of the N weighted outputs, i.e. $\hat{y} = \sum_{i=1}^{N} w_i y_i$.

decoding scheme, which we refer to as *optimal weighted decoding*. For different threshold value settings, the MSE distortion curve exhibits the SSR effect as a function of noise level and increased numbers of comparators. We compare the optimal weighted decoding scheme obtained by weighting before summation to that of weighting after summation, by analyzing the MSE distortions of each. The results show that optimal weighting of the binary quantizers' outputs before summation is superior to the case assumed in the original array, where the unweighted binary responses are simply summed. We demonstrate that optimally weighting the responses reduces the MSE distortion between the original input signal and the decoded output signal. However, we also find that there are parameter regions where optimal weighting provides a negligible reduction in mean square error, and in these regions it is therefore beneficial to avoid the additional complexity required in finding the optimal weightings and applying them.

This paper is organized as follows: Section 2 gives mathematical descriptions of optimal weighted decoding for an array of comparators. Section 3 develops the MSE distortion performance of weighted decoding for three examples of threshold setting configuration. Section 4 compares the MSE distortions between the cases of weighting before and after summation. Finally, we present the conclusions and discuss further research directions.

2. Optimal weighted decoding scheme

We here consider the weighted summing array of N noisy comparators, as shown in Fig. 1. All comparators receive the same continuously valued input signal x with standard deviation σ_x. The ith comparator is subject to independent and identically distributed (i.i.d.) additive noise components η_i with standard deviation σ_η, which are independent of the signal x. The output from each comparator, y_i, is unity if the input signal plus the noise is greater than its threshold θ_i, and zero otherwise. The noisy binary output of each individual comparator y_i is then multiplied by the weighted coefficient w_i ($w_i \in \Re$), resulting in the weighted output $w_i y_i$. All weighted outputs are summed to give the overall output $\hat{y} = \sum_{i=1}^{N} w_i y_i$.

When all weighted coefficients w_i $(i = 1, \ldots, N)$ are equal to unity, the model is identical to that studied in [22]. It is effectively a stochastic quantizer [40–43]. The summation of the outputs of all the comparators is a discretely valued stochastic encoding of x, which can take integer values between zero and N. For obtaining reconstructed signal, we need a decoding method to decode the output signal. This is performed by weighting after summation. When the weighted coefficient w_i $(i = 1, \ldots, N)$ is arbitrarily chosen, the model achieves a decoding function that is performed by weighting before summation.

2.1. Wiener linear decoding

Before considering how to optimally weight the quantizer responses, we first review what is known as *Wiener linear decoding*, as studied in [43]. In this case, we introduce y to denote the unweighted sum of the quantizer response, i.e.

$$y = \sum_{i=1}^{N} y_i. \quad (1)$$

It is shown in [43] that, under the condition where all threshold levels are identical and equal to the signal mean, and both the signal and noise have even probability density functions, that $\mathrm{E}[y] = N/2$. Under these conditions, it is of interest to consider how to optimize the MSE between the input signal, x, and a linear decoding of y written in the form

$$\hat{y}_{\mathrm{w}} = \frac{2c}{N} y - c. \quad (2)$$

The result of this operation, $\hat{y}_{\mathrm{w}}$ can be thought of as the reconstructed value of the input signal, with the error between the input x, and the reconstructed output $\hat{y}_{\mathrm{w}}$ being

$$\epsilon = x - \hat{y}_{\mathrm{w}}. \quad (3)$$

It is straightforward to derive the optimal solution for c as

$$c = \frac{N\mathrm{E}[xy]}{2\,\mathrm{var}[y]}, \quad (4)$$

where $\mathrm{var}[y] = \mathrm{E}[y^2] - N^2/4$ is the variance of y [43]. This is known as the Wiener optimal linear decoding scheme for minimizing MSE distortion [44]. The MSE distortion for Wiener decoding can be written as [43]

$$\mathrm{MSE}_{\mathrm{w}} = \mathrm{E}[x^2]\left(1 - \frac{\mathrm{E}[xy]^2}{\mathrm{E}[x^2]\,\mathrm{var}[y]}\right) = \mathrm{E}[x^2](1 - \rho_{xy}^2), \quad (5)$$

where ρ_{xy} is the correlation coefficient between the input signal x and the output y. Equation (5) also shows that the MSE distortion of Wiener decoding scheme is entirely dependant on the correlation coefficient ρ_{xy}.

2.2. Optimal weighted decoding

We now consider the case shown in Fig. 1, where arbitrary multiplicative weightings w_i $(i = 1, \ldots, N)$ are applied to the binary quantizer outputs. We seek to choose the optimal weights, $\mathbf{w}^{\mathrm{o}} = [w_1^{\mathrm{o}}, w_2^{\mathrm{o}}, \cdots, w_N^{\mathrm{o}}]^\top$ under which the MSE distortion between the decoded signal and the input is the minimum. We denote this decoding scheme as *optimal weighted decoding*, and find the optimal weights by applying least squares regression to a data obtained by simulating a sequence of samples from the input signal, and the resulting binary quantizer responses from each sample.

To begin, we introduce a vector $\mathbf{x}$ of size $(K \times 1)$ to denote a sequence of K independent samples drawn from the input signal's probability distribution. We also introduce a matrix $\mathbf{Y}$ of size $(K \times N)$ to denote the N threshold responses for each of the K input samples. We denote an arbitrary vector of weights as $\mathbf{w} = [w_1, w_2, \cdots, w_N]^\top$ and the optimal weights as $\mathbf{w}^{\mathrm{o}}$. Ideally, we desire $\mathbf{w}^{\mathrm{o}}$ to satisfy

$$\mathbf{Y}\mathbf{w}^{\mathrm{o}} = \mathbf{x}. \quad (6)$$

However, for $K > N$ (in practice we desire $K \gg N$), this is an overcomplete system of linear equations, and we therefore follow the standard approach of seeking to find $\mathbf{w}^{\mathrm{o}}$ that minimizes the MSE distortion for the data, i.e., the solution to the optimization problem

$$\mathbf{w}^{o} = \arg\min_{\mathbf{w}} \sum_{k=1}^{K} (\mathbf{y}_k \mathbf{w} - x_k)^2, \tag{7}$$

where $\mathbf{y}_k$ is the k-th row of $\mathbf{Y}$. The minimum of the sum of squares in Eq. (7) is found by setting the gradient to zero, and we have

$$\mathbf{Y}^{\top}\mathbf{x} = \mathbf{Y}^{\top}\mathbf{Y}\mathbf{w}^{o}. \tag{8}$$

When the inverse matrix $(\mathbf{Y}^{\top}\mathbf{Y})^{-1}$ exists, an exact solution to Eq. (7) can be written as

$$\mathbf{w}^{o} = (\mathbf{Y}^{\top}\mathbf{Y})^{-1}\mathbf{Y}^{\top}\mathbf{x}. \tag{9}$$

In practice, if the inverse matrix does not exist, we can regularize Eq. (9) by solving [45]

$$\mathbf{w}^{o} = (\mathbf{Y}^{\top}\mathbf{Y} + \lambda\mathbf{I})^{-1}\mathbf{Y}^{\top}\mathbf{x}, \tag{10}$$

where λ is a parameter that can be optimized using cross-validation, and $\mathbf{I}$ is the $N \times N$ identity matrix.

Having obtained the optimal weights, the resulting optimally weighted decoding $\hat{y}$ for any input sample x can be expressed as

$$\hat{y} = \sum_{i=1}^{N} w_i^{o}(y_i - \bar{y}_i), \tag{11}$$

where $\bar{y}_i$ is the mean value of each comparator output. Note that this can be expressed as

$$\hat{y} = b^{o} + \sum_{i=1}^{N} w_i^{o} y_i = b^{o} + \mathbf{y}\mathbf{w}^{o}, \tag{12}$$

where b^{o} is an optimal constant bias and $\mathbf{y}$ is the vector of all binary quantizer responses.

3. Results for optimal weighted decoding

In this section, we explore three examples of threshold settings to examine the MSE distortion performance of the optimal weighted decoding scheme, as the ratio of noise standard deviation σ_η to signal standard deviation σ_x varies. As in [22,43], we denote this ratio as $\sigma = \sigma_\eta/\sigma_x$. Since both σ_x and threshold-levels applied to x have the same units, it should be noted that our notation for threshold levels (e.g. θ) are actually the dimensionless ratio θ/σ_x. Consequently, the MSE distortion values that we calculate below have units of σ_x^2.

In this paper, except in one instance explicitly indicated, we assume that both the signal and the noise are Gaussian with zero mean.

3.1. Identical thresholds

Since we assume zero mean Gaussian signal and noise, it can be shown that when all comparators have the same threshold value $\theta_i = \theta = 0$, the mean $\bar{y}_i = 0.5$ [43]. Therefore, given a set of optimal weights, w_i^{o}, they can be assigned arbitrarily to any y_i. This is equivalent to giving all y_i the same weight, provided it is equal to the average of the weights returned by the optimization. Thus we can rewrite the decoding equation of Eq. (12) to the form

$$\hat{y} = a(y - b), \tag{13}$$

where $y = \sum_{i=1}^{N} y_i$, $a = (\sum_{i=1}^{N} w_i^{o})/N$, and $b = \sum_{i=1}^{N} \bar{y}_i$. For zero-mean noise PDFs that are even functions, and for all threshold values equal to zero, $b = \mathrm{E}[y] = N/2$ [43].

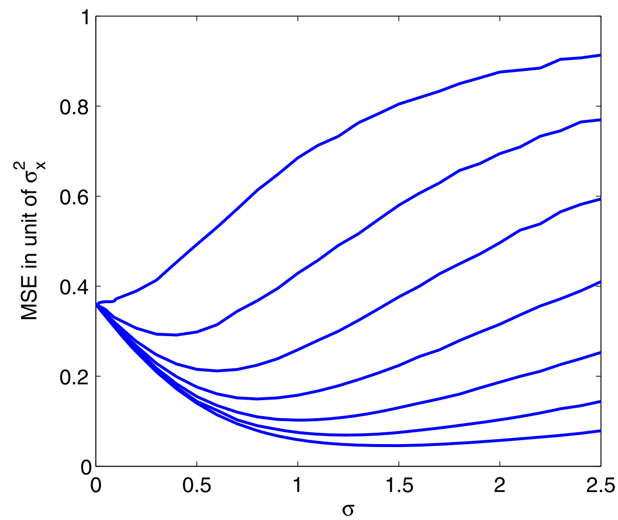

Fig. 2. The MSE distortion of weighting before summation versus σ for the case of identical threshold values, i.e. $\theta = \mathrm{E}[x] = 0$. From top to bottom, the array sizes are $N = 1, 3, 7, 15, 31, 63$ and 127.

We can write the mean square error as

$$\mathrm{MSE} = \mathrm{E}[(x - \hat{y})^2] = \mathrm{E}[\hat{y}^2] - 2\mathrm{E}[x\hat{y}] + \mathrm{E}[x^2]. \tag{14}$$

Substituting Eq. (13) into Eq. (14), the MSE distortion can be written as

$$\mathrm{MSE} = a^2\mathrm{E}[y^2] - 2a\mathrm{E}[xy] + \mathrm{E}[x^2] - \frac{a^2N^2}{4}. \tag{15}$$

When differentiating Eq. (15) with respect to a, and setting the result to zero, an optimal expression for a is

$$a = \frac{\mathrm{E}[xy]}{\mathrm{var}[y]}. \tag{16}$$

Since $b = N/2$, from Eq. (16), it is seen that Eq. (13) is actually the same as Eq. (2) when this expression for a is used. This result indicates that the optimal weighted decoding achieves the same performance as the Wiener linear decoding for the case of identical thresholds. This is also verified by the MSE distortion of Eq. (15), as shown in Fig. 2.

For all threshold values $\theta = \mathrm{E}[x] = 0$, Fig. 2 shows the MSE distortion of weighting before summation versus σ. As the number of comparators, N, increases, the MSE distortion can be optimized by tuning σ. The results are consistent with previous results obtained by Wiener linear decoding in [43]. The MSE distortions of other signal and noise distributions (not shown here) also accord with the corresponding results in [43].

Besides the identical threshold setting, it is also of interest to examine the MSE distortion of weighting before summation when each comparator has different threshold settings. In the following, we will consider two schemes of the threshold settings.

3.2. Unique thresholds

Here we consider a situation that occurs in a conventional uniform scalar quantizer, namely that the threshold levels are set uniquely at intervals of $1/(N+1)$ across the signal dynamic range. Hence, assuming a signal with a support of size σ_x, the threshold value of each comparator is given by $\theta_i = \sigma_x(i-1)/(N+1)$ for $i = 1, 2, \cdots, N$.

It is known that in the absence of noise, the best performance can be produced for the uniform quantization scheme when the input signal is also a uniform distribution [22]. Hence, uniform signal and noise distributions are used here to analyze the property

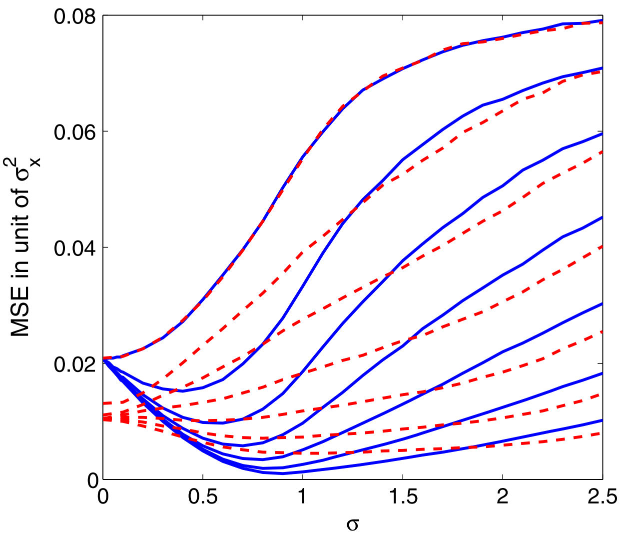

Fig. 3. The MSE distortion of weighting before summation versus the ratio of σ for uniform signal and noise distributions (from top to bottom, $N = 1, 3, 7, 15, 31, 63$, and 127). The red dashed lines show the MSE distortion for the unique threshold setting, and the blue solid lines correspond to the MSE distortion of the identical threshold setting. The optimal weights for each comparator are solved using Eq. (10) with $K = 10^6$ data samples. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of weighting before summation. The MSE distortion with unique thresholds is shown in Fig. 3 versus σ for different N (dashed red lines). Here, the optimal weights for each comparator are found by solving Eq. (10), for $K = 10^6$ data samples. For comparison, the solid blue lines correspond to the MSE distortion curves where all thresholds are set to the signal mean, as shown in Fig. 3. It is observed that, for the same size N, the case of all thresholds being identical, with respect to the case of unique thresholds, achieves a much smaller MSE distortion at the corresponding optimal nonzero value of σ. Moreover, the configuration of quantizers with the same threshold value has lower complexity than the scheme of setting N different threshold values.

3.3. Group thresholds

The third configuration of threshold setting we shall consider is dividing the set of threshold levels into two groups. In each group with size N_m $(m = 1, 2)$, all comparators have the same threshold value θ_m, while the two groups have different threshold values $\theta_1 \neq \theta_2$. Thus, the decoding output for this case of group threshold setting can be written as

$$\hat{y} = \sum_{i=1}^{N_1} w_{i,1}^{o}(y_{i,1} - \bar{y}_i) + \sum_{j=1}^{N_2} w_{j,2}^{o}(y_{j,2} - \bar{y}_j). \tag{17}$$

Within each summation on the right side of Eq. (17), the threshold values are equal, and therefore so are the means. Since there is no way to distinguish between each y_i, the weights can be assigned to any y_i. The reconstructed signal $\hat{y}$ is given by

$$\hat{y} = a_1(y_1 - b_1) + a_2(y_2 - b_2), \tag{18}$$

where the outputs of the encoder in each group are

$$y_1 = \sum_{i=1}^{N_1} y_{i,1}, \qquad y_2 = \sum_{j=1}^{N_2} y_{j,1}. \tag{19}$$

Here, the constants a_1 and a_2 can be expressed as

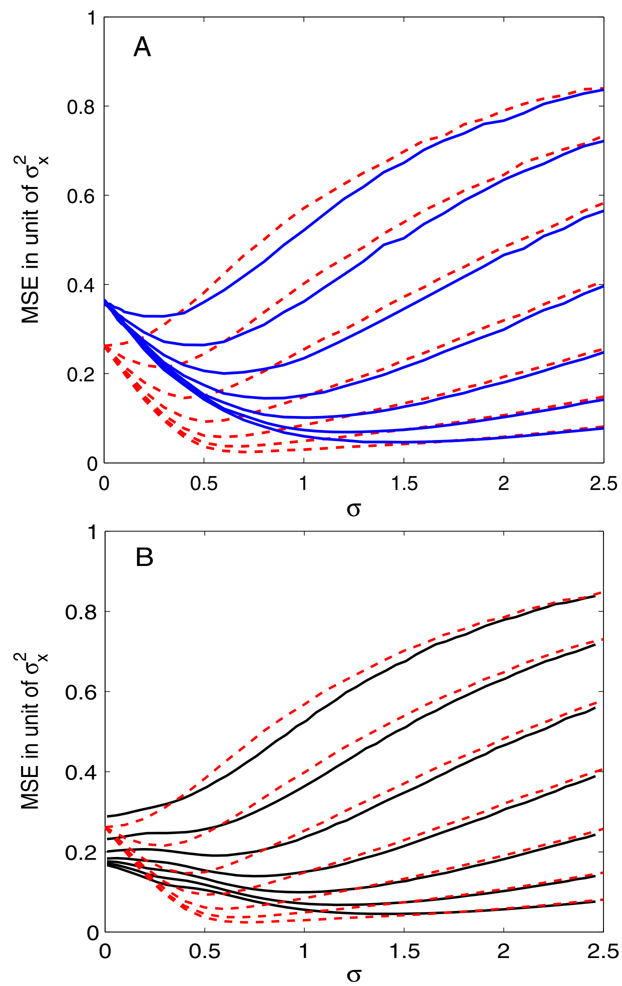

Fig. 4. The MSE distortion of weighting before summation versus σ, where from top to bottom $N = 2, 4, 8, 16, 32, 64, 128$. For comparison, (A) the red dashed lines are for the case of group thresholds, and the blue solid lines are for the case of identical thresholds; (B) The black solid lines are for the case of unique thresholds, and the red dashed lines are still for the case of group thresholds. For the group threshold setting, the quantizer size of each group is equal, that is $N_1 = N_2 = N/2$, while the threshold values are set as $\theta_1/\sigma_x = -1$ and $\theta_2/\sigma_x = 1$, respectively. The optimal weights for each comparator are found by solving Eq. (10). Data samples $K = 5 \times 10^5$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$$a_1 = (\sum_{i=1}^{N_1} w_i^{o})/N_1, \qquad a_2 = (\sum_{j=1}^{N_2} w_j^{o})/N_2, \tag{20}$$

while the constant parameters b_1 and b_2 are computed as

$$b_1 = \sum_{i=1}^{N_1} \bar{y}_i, \qquad b_2 = \sum_{j=1}^{N_2} \bar{y}_j. \tag{21}$$

The MSE distortions are plotted versus σ for different array sizes (dashed red lines) in Fig. 4(A). Here, the optimal weights for each comparator are found by solving Eq. (10) with $K = 5 \times 10^5$ data samples. Two groups are with the same size $N_1 = N_2 = N/2$, but different threshold values of $\theta_1/\sigma_x = -1$ and $\theta_2/\sigma_x = 1$. For comparison, the solid lines represent the MSE distortions for the case of the identical threshold value setting. It is obvious that, for large N, the optimal MSE distortion for the case of group threshold levels is much lower than that of identical threshold levels. A careful observation shows that, for the case of group thresholds, the position of the minimum MSE distortion shifts to lower noise intensities, rather than higher noise intensities as seen for identical thresholds. More interestingly, for the group threshold setting,

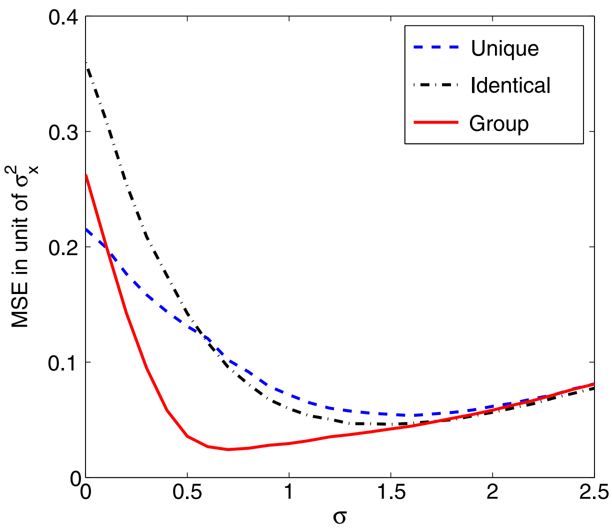

Fig. 5. The MSE distortion of weighting before summation against increasing σ when the threshold values are identical, unique and grouped. Here, the number of quantizers is $N = 128$.

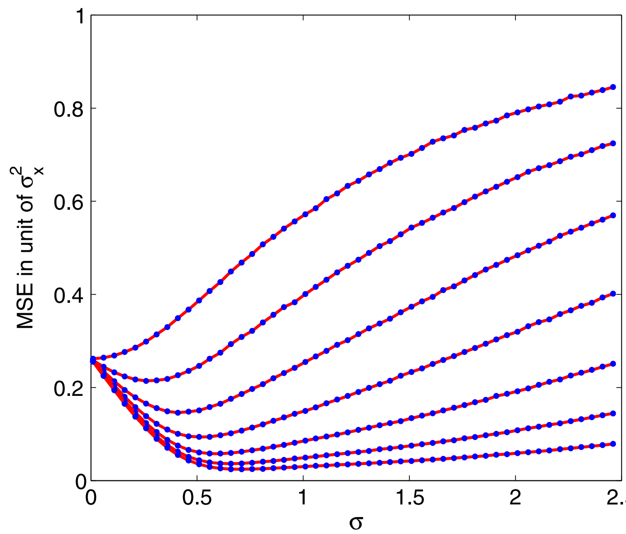

Fig. 6. The MSE distortion of weighting before summation compared to that of weighting after summation against increasing σ in the case of group thresholds. Here, from top to bottom, the curves of MSE distortion are plotted for $N = 2, 4, 8, 16, 32, 64, 128$. The sizes of each group are $N_1 = N_2 = N/2$, and the threshold values within each group are $\theta_1/\sigma_x = +1$ and $\theta_2/\sigma_x = -1$, respectively. The red solid lines show the MSE distortions of weighting before summation, and the blue circles lines correspond to that of weighting after summation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the MSE shows the monotonic increasing behavior for $N = 2$. For $N > 2$ quantizers, the MSE distortion can be reduced by optimally tuning σ. Similarly, we also compare the case of group thresholds with the case of unique thresholds (black solid lines) in Fig. 4(B). It is shown that, for large N, the MSE distortion for the case of unique thresholds is much larger than that of group thresholds. Other signal and noise distributions also demonstrate similar results (not shown here for simplicity).

3.4. Comparison of performance for different threshold setting strategies

We now directly compare the MSE distortion for the optimal weighted decoding method across each of the three cases of identical, unique and group threshold settings. Fig. 5 shows the MSE distortion for the above three cases for $N = 128$. It is seen in Fig. 5 that the group threshold case gives the smallest MSE distortion, and possesses the best performance for almost the whole range of σ. A further observation is that the minimum MSE for the group threshold values occurs at an optimal level of $\sigma < 1$, while the identical and unique threshold cases have slightly higher minima at a much larger σ ($\sigma > 1$).

At $\sigma = 0$, the MSE of the identical thresholds is the largest among the three cases, while the MSE of the unique thresholds is the smallest. The reason is that, for the identical thresholds, all thresholds switch in unison, only a binary output signal with state 0 or 1 can be obtained, and only a binary representation of the input signal is available, resulting in a large MSE distortion. In comparison, for the case of N unique thresholds (which is a special case of the group threshold setting), the output is an $N + 1$ state representation of the input signal, which enables a smaller MSE distortion.

These comparisons motivated us to seek a method for decreasing the MSE distortion using the group threshold setting. As σ increases, it is seen in Fig. 5 that the MSE of the two group threshold setting is much better than other threshold settings. It is also, however, shown in Fig. 5 that, for very large noise levels, the MSE distortions for all threshold settings tend to be equal, and the group threshold values reduce to the signal mean, i.e. the identical thresholds case. These results are expected, because previous work has shown that group thresholds are optimal except for very small noise levels and for large noise; in the former case, unique thresholds are optimal, while in the latter case, identical thresholds are optimal [42,43]—our results in this paper are consistent with this prior work. The task of determining the optimal number of group threshold levels, group sizes and threshold values is an interesting question for future work.

4. MSE distortion comparison between the case of weighting before and after summation

In Section 3, it is observed that the performance of weighted decoding in the case of a group threshold setting is better than other threshold setting schemes. Thus, it is interesting to compare MSE distortions between optimal weighted decoding versus Wiener linear decoding; that is the MSE distortion comparison between the cases of weighting before and after summation. It is also known that the expected value of the encoding output is $E[y] = N/2$ for the case of identical thresholds, but this result will not necessarily hold for the case of group thresholds. Under such a condition, the reconstruction points using the Wiener decoding technique are different from Eq. (2), and take on the following form [43]

$$\hat{y} = \frac{E[xy]}{\mathrm{var}[y]}(y - E[y]), \quad \text{for } y = 0, 1, \cdots, N. \tag{22}$$

Fig. 6 shows the comparison of performance between the cases of weighting before and after summation with group thresholds. We here choose sizes of each group $N_1 = N_2 = N/2$, and the threshold values for each group as $\theta_1/\sigma_x = +1$ and $\theta_2/\sigma_x = -1$, respectively. In Fig. 6, the red solid and blue circles lines represent the MSE distortions of weighting before and after summation, respectively. The results demonstrate that weighting before summation gives almost the same performance as weighting after summation. The reason is that if the threshold values θ_1 and θ_2 are mutually opposite real numbers, as illustrated in Fig. 6, a_1 is then equal to a_2 in Eq. (18).

In order to further identify the MSE distortion difference between weighting before and after summation in the case of group thresholds, we define the percentage P as

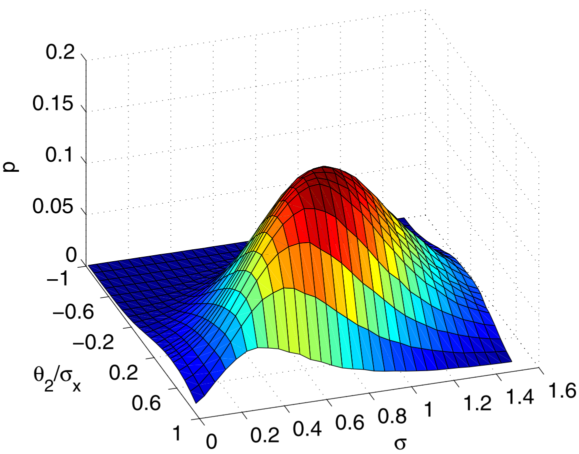

Fig. 7. The MSE distortion percentage difference between weighting before and after summation against the ratio σ and the thresholds θ_2/σ_x, for the given group thresholds $\theta_1/\sigma_x = 1$ and the total quantizer size $N = 128$. The array sizes of each group are $N_1 = N_2 = N/2$, respectively.

$$P = \frac{\text{MSE}_{\text{after}} - \text{MSE}_{\text{before}}}{\text{MSE}_{\text{before}}}, \tag{23}$$

where $\text{MSE}_{\text{after}}$, $\text{MSE}_{\text{before}}$ are the distortion values for weighting after and before summation, respectively. In this way, for the given group threshold level $\theta_1/\sigma_x = 1$ and array size $N = 128$, the difference percentage P between the two decoding schemes, for various θ_2/σ_x and σ, is shown in Fig. 7.

In Fig. 7, it is seen that, for the symmetrical setting of $\theta_1/\sigma_x = 1$ and $\theta_2/\sigma_x = -1$, the percentage difference is extremely small, but always positive. This means that weighting after summation is always slightly worse than weighting before summation. However, it is also clear that in this case, summing and carrying out Wiener linear decoding is nearly as good as optimal weighted decoding, because the distortion is only negligibly reduced. Consequently, under these conditions the additional complexity required to obtain optimal weights and use them in an array is unlikely to be warranted, and the summing without first weighting used in the original array can be considered to be a superior array configuration.

However, it is also seen that, as the group threshold value θ_2/σ_x increases, the difference percentage P presents a maximum peak at $\theta_2/\sigma_x = 0.35$ and $\sigma = 0.7$. For the group threshold levels of $\theta_1/\sigma_x = 1$ and $\theta_2/\sigma_x = 0.35$, the MSE distortion curves of two decoding schemes are illustrated in Fig. 8(A). It illustrates that weighting before summation gives rise to a smaller MSE. Moreover, as N increases, the difference between the MSE distortions for the two decoding schemes becomes larger. The percentage differences P between two decoding methods are also clearly exhibited in Fig. 8(B) for group sizes $N_1 = N_2 = N/2$. It demonstrates that weighting before summation is better than the case of weighting after summation. Therefore, the conclusion that summing before carrying out Wiener decoding does not significantly lower performance is very much dependent on the choice of thresholds.

5. Conclusion and discussions

In this paper, we study the problem of optimizing the decoding output signal through a summing array of binary quantizers. A new *optimal weighted decoding* scheme is proposed, and the MSE distortions of this decoding scheme for three examples of threshold settings, i.e., identical, unique and group thresholds, are analyzed in detail. The obtained results show that the MSE distortion is the lowest for the case of group thresholds. In this situation, the MSE distortion of optimal weighted decoding, i.e. performed by weighting before summation, is improved over that of weighting after summation. In addition, decoding performance of weighting after summation is nearly as good as that of weighting before summation in certain threshold groups. This is a significant finding for many potential applications inspired by SSR, because summing without first weighting is much simpler than weighting and then summing, and its MSE distortion performance is not reduced.

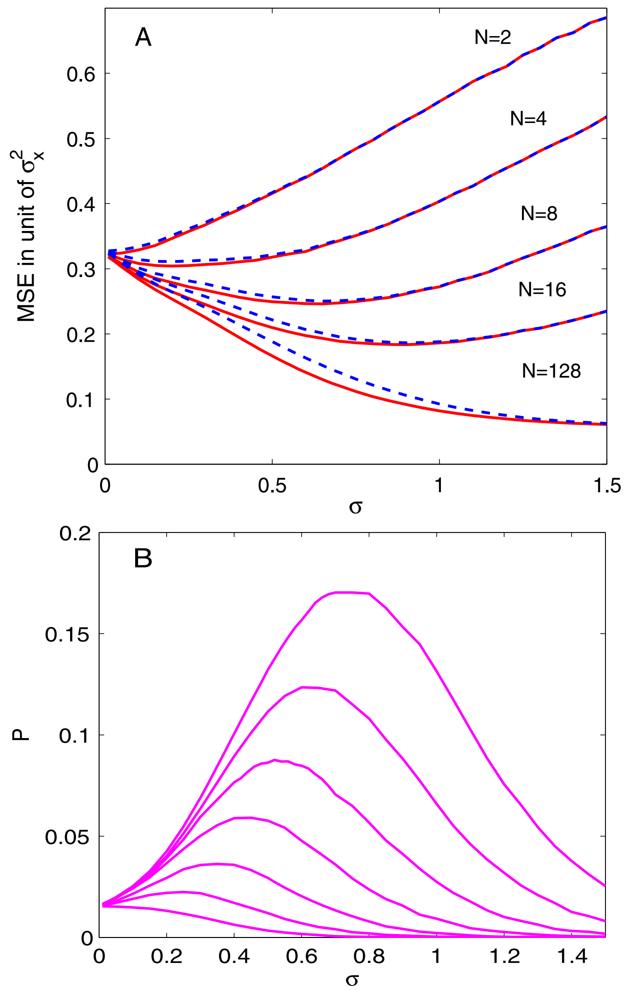

Fig. 8. (A) The MSE distortion of weighting before summation compared to that of weighting after summation against increasing σ for the case of group thresholds, from top to bottom $N = 2, 4, 8, 16, 128$. The red solid lines show the results of weighting before summation, and the blue dashed lines show the results of weighting after summation. (B) The MSE distortion percentage difference between weighting before and after summation against increasing σ for the case of group thresholds, from bottom to top $N = 2, 4, 8, 16, 32, 64, 128$. The array sizes of each group are $N_1 = N_2 = N/2$, respectively. The threshold values within each group are $\theta_1/\sigma_x = 1$ and $\theta_2/\sigma_x = 0.35$, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

We note that, from the viewpoint of quantization theory, the MSE distortion describes the average error between the input signal and the reconstructed output. The initial SSR research mainly considers the viewpoint of information transmission, and thus the mutual information is frequently used to analyze the channel capacity of SSR model. It will be very interesting to find the inherent relationship between the two measures in the same situation; for example, whether the optimal noise levels that correspond to the minimum MSE or the maximum mutual information are the same or not.

From the point of view of signal estimation theory, the minimum MSE distortion will be no less than a theoretical limit given by the average information bound [46], which is related to the bias and the Fisher information of an estimator (i.e. the output of the SSR model). Naturally, it is of future interest to develop a mathematical description for the minimum MSE distortion that the optimal weighted decoding approach can achieve. This is an unconstrained multidimensional optimization problem, since the group sizes, the threshold value settings and the noise level will be all unconstrained variables of the MSE distortion. How much lower can the minimum MSE distortion of weighting before summation be than that of weighting after summation and to what extent does it approach the theoretical average information bound?

Previously, mathematical expressions for the large N Wiener linear decoding MSE distortion, and exact results for specific signal and noise distributions have been derived [43]. For the optimal weighted decoding, with $N_{1,2}$ going to infinity, the MSE expression is more complicated than that of the optimal Wiener decoding. Thus, there remains an open question that is of interest for our future work, which is whether MSE distortion has asymptotic behavior or not.

Acknowledgements

This work is sponsored by the Science and Technology Development Program of Shandong Province (No. 2014GGX101031). Mark D. McDonnell's contribution was supported by an Australian Research Fellowship from the Australian Research Council (project number DP1093425).

References

[1] R. Benzi, A. Sutera, A. Vulpiani, The mechanism of stochastic resonance, J. Phys. A, Math. Gen. 14 (1981) L453–L457.
[2] L. Gammaitoni, P. Hänggi, P. Jung, F. Marchesoni, Stochastic resonance, Rev. Mod. Phys. 70 (1998) 223–287.
[3] M.I. Dykman, P.V.E. McClintock, What can stochastic resonance do?, Nature 391 (1998) 344.
[4] S. Kay, Can detectability be improved by adding noise?, IEEE Signal Process. Lett. 7 (2000) 8–10.
[5] D. Abbott, Overview: unsolved problems of noise and fluctuations, Chaos 11 (2001) 526–538.
[6] S. Fauve, F. Heslot, Stochastic resonance in a bistable system, Phys. Lett. A 97 (1983) 5–7.
[7] B. McNamara, K. Kiesenfeld, R. Roy, Observation of stochastic resonance in a ring laser, Phys. Rev. Lett. 60 (1998) 2626–2629.
[8] A. Longtin, Stochastic resonance in neuron models, J. Stat. Phys. 70 (1993) 309–327.
[9] K. Wiesenfeld, F. Moss, Stochastic resonance and the benefits of noise: from ice ages to crayfish and SQUIDs, Nature 373 (1995) 33–36.
[10] S.M. Bezrukov, I. Vodyanoy, Noise-induced enhancement of signal transduction across voltage dependent ion channels, Nature 378 (1995) 362–364.
[11] J.J. Collins, C.C. Chow, T.T. Imhoff, Aperiodic stochastic resonance in excitable systems, Phys. Rev. E 52 (1995) R3321–R3324.
[12] A.R. Bulsara, A. Zador, Threshold detection of wideband signals: a noise-induced maximum in the mutual information, Phys. Rev. E 54 (1996) R2185–R2188.
[13] C. Heneghan, C.C. Chow, J.J. Collins, T.T. Imhoff, S.B. Lowen, M.C. Teich, Information measures quantifying aperiodic stochastic resonance, Phys. Rev. E 54 (3) (1996) R2228–R2231.
[14] A. Neiman, D. Shulgin, V. Anishchenko, W. Ebeling, L. Schimansky-Geier, J. Freund, Dynamical entropies applied to stochastic resonance, Phys. Rev. Lett. 76 (23) (1996) 4299–4302.
[15] F. Chapeau-Blondeau, Noise-enhanced capacity via stochastic resonance in an asymmetric binary channel, Phys. Rev. E 55 (2) (1997) 2016–2019.
[16] A. Capurro, K. Pakdaman, T. Nomura, S. Sato, Aperiodic stochastic resonance with correlated noise, Phys. Rev. E 58 (1998) 4820–4827.
[17] F. Chapeau-Blondeau, D. Rousseau, Noise improvements in stochastic resonance: from signal amplification to optimal detection, Fluct. Noise Lett. 2 (2002) L221–L233.
[18] G.P. Harmer, B.R. Davies, D. Abbott, A review of stochastic resonance: circuits and measurement, IEEE Trans. Instrum. Meas. 51 (2002) 299–309.
[19] A. Patel, B. Kosko, Stochastic resonance in noisy spiking retinal and sensory neuron models, Neural Netw. 18 (2005) 467–478.
[20] F. Duan, F. Chapeau-Blondeau, D. Abbott, Fisher-information condition for enhanced signal detection via stochastic resonance, Phys. Rev. E 84 (2011) 051107.
[21] F. Duan, F. Chapeau-Blondeau, D. Abbott, Fisher information as a metric of locally optimal processing and stochastic resonance, PLoS ONE 7 (2012) e34282.
[22] N.G. Stocks, Suprathreshold stochastic resonance in multilevel threshold systems, Phys. Rev. Lett. 84 (2000) 2310–2313.
[23] N.G. Stocks, Information transmission in parallel arrays of threshold elements: suprathreshold stochastic resonance, Phys. Rev. E 63 (2001) 041114.
[24] N.G. Stocks, Suprathreshold stochastic resonance: an exact result for uniformly distributed signal and noise, Phys. Lett. A 279 (2001) 308–312.
[25] N.G. Stocks, R. Mannella, Generic noise enhanced coding in neuronal arrays, Phys. Rev. E 64 (2001) 030902(R).
[26] N.G. Stocks, D. Allingham, R.P. Morse, The application of suprathreshold stochastic resonance to cochlear implant coding, Fluct. Noise Lett. 2 (3) (2002) L169–L181.
[27] H. Li, Y. Wang, Noise-enhanced information transmission of a non-linear multilevel threshold neural networks system, Acta Phys. Sin. 63 (12) (2014) 120506.
[28] V.N. Hari, G.V. Anand, A.B. Premkumar, A.S. Madhukumar, Design and performance analysis of a signal detector based on suprathreshold stochastic resonance, Signal Process. 92 (7) (2012) 1745–1757.
[29] Y. Guo, J. Tan, Suprathreshold stochastic resonance of a non-linear multilevel threshold neuronal networks system, Acta Phys. Sin. 61 (17) (2012) 170502.
[30] Y. Guo, J. Tan, Suprathreshold stochastic resonance in multilevel threshold system driven by multiplicative and additive noises, Commun. Nonlinear Sci. Numer. Simul. 18 (2013) 2852–2858.
[31] B. Zhou, M.D. McDonnell, Optimising threshold levels for information transmission in binary threshold networks: independent multiplicative noise on each threshold, Physica A 419 (2015) 659–667.
[32] S. Zozor, P.O. Amblard, C. Duchene, On pooling networks and fluctuation in suboptimal detection framework, Fluct. Noise Lett. 7 (1) (2007) L39–L60.
[33] M.D. McDonnell, P.O. Amblard, N.G. Stocks, Stochastic pooling networks, J. Stat. Mech. Theory Exp. 1 (2009) p01012.
[34] R. Mathar, A. Schmeink, Cooperative detection over multiple parallel channels: a principle inspired by nature, in: IEEE 22nd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), 2011, pp. 1758–1762.
[35] M.D. McDonnell, D. Abbott, C.E.M. Pearce, A characterization of suprathreshold stochastic resonance in an array of comparators by correlation coefficient, Fluct. Noise Lett. 2 (2002) L205–L220.
[36] D. Rousseau, F. Duan, F. Chapeau-Blondeau, Suprathreshold stochastic resonance and noise-enhanced Fisher information in arrays of threshold devices, Phys. Rev. E 68 (2003) 031107.
[37] D. Rousseau, F. Chapeau-Blondeau, Suprathreshold stochastic resonance and signal-to-noise ratio improvement in arrays of comparators, Phys. Lett. A 321 (2004) 280–290.
[38] S. Durrant, Y. Kang, N.G. Stocks, J. Feng, Suprathreshold stochastic resonance in neural processing tuned by correlation, Phys. Rev. E 84 (2011) 011923.
[39] P.E. Greenwood, P. Lansky, Information content in threshold data with non-Gaussian noise, Fluct. Noise Lett. 7 (2007) L79–L89.
[40] M.D. McDonnell, D. Abbott, C.E.M. Pearce, An analysis of noise enhanced information transmission in an array of comparators, Microelectron. J. 33 (2002) 1079–1089.
[41] M.D. McDonnell, N.G. Stocks, C.E.M. Pearce, D. Abbott, Quantization in the presence of large amplitude threshold noise, Fluct. Noise Lett. 5 (2005) L457–L468.
[42] M.D. McDonnell, N.G. Stocks, C.E.M. Pearce, D. Abbott, Optimal information transmission in nonlinear arrays through suprathreshold stochastic resonance, Phys. Lett. A 352 (2006) 183–189.
[43] M.D. McDonnell, N.G. Stocks, C.E.M. Pearce, D. Abbott, Stochastic Resonance: From Suprathreshold Stochastic Resonance to Stochastic Signal Quantization, Cambridge University Press, 2008.
[44] R.D. Yates, D.J. Goodman, Probability and Stochastic Processes: A Friendly Introduction for Electrical and Computer Engineers, second edn., John Wiley and Sons, 2005.
[45] G.-B. Huang, H. Zhou, X. Ding, R. Zhang, Extreme learning machine for regression and multiclass classification, IEEE Trans. Syst. Man Cybern., Part B, Cybern. 42 (2012) 513–529.
[46] T.M. Cover, J.A. Thomas, Elements of Information Theory, John Wiley and Sons, 1991.