CHAOS SHIFT KEYING IN THE PRESENCE OF NOISE: A SIMPLE DISCRETE TIME EXAMPLE

Martin Hasler
Department of Electrical Engineering
Swiss Federal Institute of Technology
1015 Lausanne, Switzerland

ABSTRACT
The transmission of binary information using chaotic switching, also called chaos shift keying is used. On the receiver side the information is extracted using a method that is not based on synchronization, but that directly tries to detect which chaotic system of the two can have produced the received signal, provided the channel noise is bounded by a known value. The results obtained are not yet as good as for conventional systems, but they constitute a step forward in the performance of transmission systems based on chaos. The more traditional decoding method based on chaos synchronization is also analyzed, and if used in an efficient way, also gives quite good results.

1. INTRODUCTION
Since the beginning of research on transmission of information using chaos [1] various modulation schemes have been (cf. [2] for an overview). While in the beginning the motivation was mainly to hide the information in chaos, the attention has now shifted more towards the "spread spectrum" nature of the transmission. In order to compete with the well-developed conventional spread spectrum communication techniques, however, the performance of the chaotic communication systems in the presence of noise has to be improved. The purpose of this paper is to show two methods that, in the authors opinion, have a potential to perform well in the presence of a noisy transmission channel.

2. COMMUNICATION SYSTEM
We consider the iteration of the skew tent map \( f \), centered on 0 (Fig.1). Since the only state variable of this dynamical system is directly transmitted, there is no privacy of the message. Security of the transmission is not the aim of this work.

Among the methods to provide the chaotic signal with information, chaotic switching, or chaos shift keying [2-5] is thought to be robust against noise. In this method, two similar chaos generators are used to transmit a binary information signal, one for transmitting the bit "1" and the other for transmitting the bit "0". In our case this is the iteration of \( f \) and the iteration of \(-f\). For each bit, a chaotic waveform of finite length \( N \) produced by the corresponding system is sent. The receiver then has to decide for each noise corrupted incoming waveform, which system has been used to produce it. The problem how to delimit in the incoming signal the sections corresponding to the different bits is not addressed in this paper. In the spread spectrum communications literature \( N \) is called the spreading factor.

Usually, the decoding of each waveform of length \( N \) in the receiver is performed by trying to synchronize copies of both chaos generators with the active chaos generator in the transmitter, via the transmitted signal [2-5]. The attempt should be successful for the generator that matches, and unsuccessful for the other generator. This allows to decide for the correct bit. Our second methods
follows this path. In fact, synchronization is surprisingly useful even in the presence of relatively large noise, provided that the synchronization errors are used in an adequate manner to reach the decision between the bit values. The first method is linked to a method to extract a chaotic signal from noise. Various methods to achieve this goal have been proposed in the literature [6-10]. In principle, all can be used to decode chaotic waveforms contaminated by noise. However, many are computationally intensive and thus may not be suitable for real-time decoding in the receiver. Our method is based on [10].

3. METHOD BASED ON RECOVERING TRAJECTORIES

This method is based on a very simple idea. Let $y(k), k = 1, \ldots, N$ be the received signal. Suppose we know that it is contaminated by noise of amplitude limited by $b$, but otherwise of unspecified nature. Then we apply the following algorithm.

**Step 1:** The clean signal $x$ satisfies

$$x(k) \in B_k^1 = \left\{ z \mid z(k) - y(k) < b \right\} \cap [-1, +1]$$

**Step 2:** Under the hypothesis that clean signal is produced by the iteration of $f$, we must have

$$x(k) \in B_k^2 = B_k^1 \cap f^{-1}(B_{k+1}^1)$$

**Step 3 to N:** Continue analogously, computing at step $i$ the sets $B_k^i$, $k = 1, \ldots, N-i+1$, unless the algorithm stops before.

**Early stopping criterion:** Stop if any set $B_k^i$ is empty.

**Conclusion for the first part:** If any $B_k^i$ is empty, the hypothesis is wrong. Otherwise interpret $|B_k^N|/2$ as the probability that the hypothesis was correct.

**Repeat step 2 to N with -f instead of f.**

**Conclusion for the second part:** If any $B_k^i$ is empty, the hypotheses that the clean signal was produced by the iterations of -f is wrong. Otherwise interpret $|B_k^N|/2$ as the probability that this hypothesis was correct.

**Final decision:** If the first part or the second part of the algorithm was stopped early, the decision whether $f$ or $-f$ has produced the clean signal is clear. Otherwise, i.e. when the outcome is ambiguous, decide for the function with the higher probability, i.e. the larger $|B_k^N|$.

**Discussion:**

Since $f$ and $-f$ are noninvertible maps, the inverse image of an interval is in general a union of intervals, in our case two intervals. Even though the total length of this union is equal to the length of the original interval, the two constituent intervals are smaller. In the intersection (2) often only one of them contributes and thus it is expected that at each iteration $i$ the sets $B_k^i$, which are in general a union of intervals, have a rapidly decreasing size. Furthermore, in the case of the wrong hypothesis, it is expected that rapidly an empty intersection is reached.

Numerical simulations as well as some reasoning indeed confirm this picture as long as the noise amplitude is not too large. In fact, for small noise and a suitable choice of the breakpoint $a$ of $f$, it can be seen that there is no ambiguity. If the noise is too large, then clearly the $B_k^i$ are the whole interval $[-1, 1]$ and thus nothing can be deduced from this method.

**Simulation results:**

We have explored with a number of simulations what we think is (without further optimization) about the limit of the method. The choice of the breakpoint $a$ was somewhat arbitrarily 0.27. We have taken 20 evenly spaced initial conditions and for each initial condition 1000 trials. We counted the number of wrong decisions and divided this number by the total number of trials, which gives the bit error rate (BER). These errors come from ambiguous outcomes of the algorithm. In the following table, the BER is given as a function of $N$ and the signal to noise ratio.

<table>
<thead>
<tr>
<th>SNR \ N</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>14 dB</td>
<td>0.008</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1. Bit error rate as a function of N and of the signal to noise ratio for uniform noise

The usual representation of the BER versus (Energy/bit) / (noise spectral density) is given in Fig. 2.

Inherent in the method is the assumption of a bounded noise amplitude. We have also carried out some simulations with a gaussian white noise. The algorithm was applied as if the noise was bounded by some value. In this case the outcome of the algorithm may be contradictory, indicating that both f and -f could not have produced the received signal. In this case we chose the received bit randomly.

Table 2. Bit error rate as a function of N and the signal to noise ratio for gaussian noise

4. METHOD BASED ON SYNCHRONIZATION

The receiver is consists of two dynamical systems, the iteration of f and the iteration of -f, who try to synchronize with the incoming signal y using nonlinear feedback according to the equation (for f)

\[ x(k+1) = f((1-\varepsilon)x(k) + \varepsilon y(k)) \]  (3)

It has been shown in [11] that this form of feedback has better synchronization behavior than linear feedback, and that for

\[ \varepsilon \in \left[ \frac{1+a}{2}, \frac{3-a}{2} \right] \]  (4)

synchronization is achieved with a clean input signal y, supposing a > 0. Synchronization means here

\[ e(k) = |x(k) - y(k)| \to 0 \text{ as } k \to \infty \]  (5)

Actually, in the case of a clean signal the best value of \( \varepsilon \) is 1 which trivially leads to synchronization in 1 step. In the presence of noise, however, values of \( \varepsilon \) different from 1, but lying in the interval (4) are more efficient. For a 30% noise level with uniform distribution, a typical form of \( e(k) \) is given in Fig. 3. At first sight it looks not very encouraging.

However, if we compare it to the synchronization error for the (wrong) system generated by -f (Fig. 4, note the different scale on the vertical axis, as compared to Fig. 3), it becomes clear that the two hypotheses, i.e. signal produced by f and signal produced by -f can be distinguished.
Guided by these figures, we introduced the following algorithm for the decision which bit was sent. Let $e_+(k)$ and $e_-(k)$ be the synchronization error for the iterations of $f$ and $-f$, respectively. Let

$$E_+ = \sum_{k=2}^{N} e_+(k), \quad E_- = \sum_{k=2}^{N} e_-(k), \quad (6)$$

Then we decide for bit "1", if $E_+ < E_-$ and for bit "0" if $E_+ > E_-$. 

**Simulation results:**

With $a = 0.27$, $\varepsilon = 0.85$ and the noise amplitude of 30% (10.5dB SNR), we estimate the bit error rate for $N = 6$ at BER $\approx 0.001$, for 40% noise (8dB SNR) and $N = 9$ at BER $\approx 0.004$ and for 50% noise (6dB SNR) and $N = 15$ at BER $\approx 0.007$. The computational burden grows only linearly with $N$, and it is lower even for small $N$ than when applying the first method.

5. CONCLUSION

We have given two methods that are able to extract binary information from a chaotic signal that is corrupted by noise of about $8-10$ dB SNR with a low computational effort, when chaos shift keying is used with the skew tent map as a chaos generator. The bit error rate in this case is about $10^{-3}$, but it can still be lowered by increasing the spreading factor (4-6 for the first method and 6-9 for the second method) and accordingly the amount of computation.

The method based on trajectory recovering works perfectly for smaller, amplitude bounded noise and breaks down quickly for higher noise, whereas the method based on synchronization has a more graceful degradation towards higher noise values. The computational effort per bit grows at least with $N^2$ and at most as $2^N$ for the first method and with $N$ for the second, where $N$ is the spreading factor.

It should be stressed that for the results obtained not all parameters have been optimized.

**Acknowledgements:**

Thanks are due to A.N.Sharkovsky for enlightening discussions. This work has been financially supported by the Swiss National Science Foundation, grant 7UKPJ 048229 (cooperation with the CEEC/NIS states, financed by the ministry of foreign affairs) and grant 2000-047172.96.

6. REFERENCES


