

# Two Aspects of Optimum CSK Communication: Spreading and Decoding

Theodore Papamarkou

University of Warwick  
Department of Statistics  
Coventry, CV4 7AL, UK  
(e-mail: T.Papamarkou@warwick.ac.uk)

**Abstract.** This paper focuses on the optimization of performance of single-user chaos shift-keying (CSK). More efficient signal transmission is achieved in the coherent case by introducing the class of the so called deformed circular maps for the generation of spreading. Also, the paired Bernoulli circular spreading (PBCS) is introduced as an optimal choice, which attains the lower bound of bit error rate (BER). As interest shifts to the non-coherent version of the system, attention moves to the receiver end. Maximum likelihood (ML) decoding is utilized serving as an improvement over the correlation decoder. To make the methodology numerically realizable, a Monte Carlo likelihood approach is employed.

**Keywords:** Communication systems, Chaos shift-keying, Optimal spreading, Correlation decoding, Likelihood decoding, Monte Carlo likelihood.

## 1 Single-user coherent CSK: the model

A brief description of single-user coherent CSK, based on [4], is provided.

Assume that the information source comprises of  $K$  binary bits, each of which is transmitted  $N$  times for reliability purposes. Let  $b = \pm 1$  denote any single binary bit out of the  $K$ , which is replicated  $N$  times.

A stationary process  $\mathbf{X} := (X_0, X_1, \dots, X_{N-1})$  of mean  $\mu_X$  and variance  $\sigma_X^2$  is involved in the modulation process.  $\mathbf{X}$  is called the *spreading* and  $N$  the *spreading factor*.

The transmitter emits the scalar product  $\mathbf{T} := b(\mathbf{X} - \mu_X \mathbf{1})$ .

The  $N$ -length *transmitted signal*  $\mathbf{T}$  is degraded as it passes through the channel. Stochastic *channel noise*  $\boldsymbol{\epsilon} := (\epsilon_0, \epsilon_1, \dots, \epsilon_{N-1})$  models the corruption of  $\mathbf{T}$ . It is assumed that the system is affected by *white channel noise*, which means that  $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma_\epsilon^2 \mathbf{I})$ , where  $\mathbf{0}$  and  $\mathbf{I}$  stand for the  $N$ -length null vector and the  $N \times N$  identity matrix respectively.

A first modelling approach would be to set the *received signal*  $\mathbf{R}$  to be the transmitted signal  $\mathbf{T}$  distorted by the additive channel noise  $\boldsymbol{\epsilon}$ , that is

$$\mathbf{R} := \mathbf{T} + \boldsymbol{\epsilon} = b(\mathbf{X} - \mu_X \mathbf{1}) + \boldsymbol{\epsilon}. \quad (1)$$

## 2 Criteria for optimal spreading

When evaluating the reliability of the system, simulations of bit error rate (BER) against signal-to-noise ratio (SNR) are run. BER is defined as the probability of erroneously decoding a single bit  $b$ . A lower bound has been found for BER. For more details, see [4].

Vital for the outcome of BER simulations is the *bit energy*  $S(\mathbf{X})$  of spreading  $\mathbf{X}$ , which is defined to be the function  $S(\mathbf{X}) := \|\mathbf{X} - \mu_x \mathbf{1}\|^2 / \sigma_x^2$ .

The choice of  $\mathbf{X}$  affects BER. There thus arises the question how could  $\mathbf{X}$  be selected in order to minimize BER. In an attempt to offer an answer, the concept of bit energy helped to form three criteria for optimally choosing the spreading.

They are summarized here hierarchically from the most to the least stringent. As the firmness of the condition imposed by each of the three successive criteria loosens, the possibility of attaining the condition increases with a relative payoff in the system's performance.

**Optimality Criterion I:** Define the spreading  $\mathbf{X}$  in a way that its bit energy  $S(\mathbf{X})$  becomes a constant equal to the spreading length  $N$ .  $\mathbf{X}$  is then optimal.

**Optimality Criterion II:** Set  $\mathbf{X}$  such that  $S(\mathbf{X})$  has zero variance.

**Optimality Criterion III:** Choose  $\mathbf{X}$  whose mean-adjusted quadratic autocorrelation has  $lag(1)$  which is as close as possible to  $-1$ , that is

$$Corr\left[\left(X_t - \mu_x\right)^2, \left(X_{t+1} - \mu_x\right)^2\right] = -1. \quad (2)$$

[8] and [5] elaborate on how these criteria have been derived.

## 3 Family of deformed circular maps

One way of generating the stationary spreading  $\mathbf{X}$  would be to start from a random variable  $X_0$  and iteratively produce the rest of the  $N - 1$  members of  $\mathbf{X}$  by means of an one-dimensional map  $\tau : [c, d] \rightarrow [c, d]$ ,  $c, d \in \mathbb{R}$ . Then,

$$X_n = \tau(X_{n-1}), \quad n = 1, 2, \dots, N - 1. \quad (3)$$

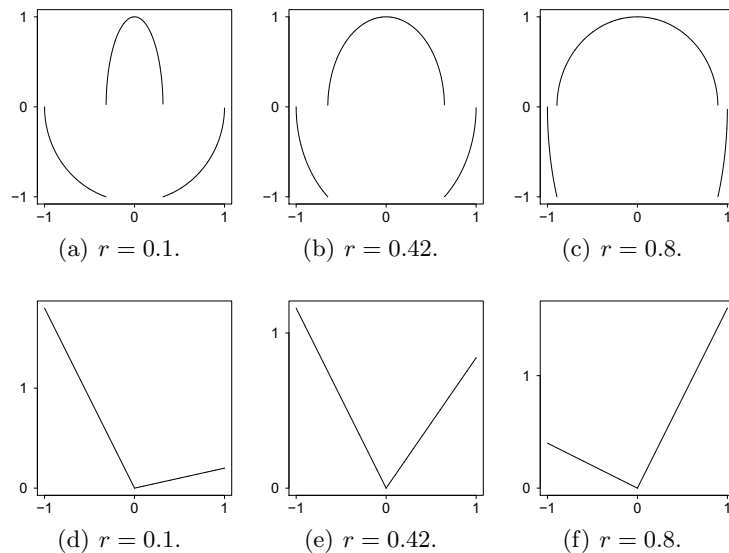
Widely used choices of  $\tau$  are, as instances, the tent, Bernoulli and logistic maps. However,  $\tau$  can be defined in ways that better conform with the third optimality criterion.

In trying to reduce the  $lag(1)$  of mean-adjusted quadratic autocorrelation function (ACF), the class of *deformed circular maps* has been defined as

$$\tau(x) := \begin{cases} -\sqrt{-(1-r)^{-1}x^2 + (1-r)^{-1}}, & -1 \leq x < -\sqrt{r} \\ \sqrt{-r^{-1}x^2 + 1}, & -\sqrt{r} \leq x < \sqrt{r} \\ -\sqrt{-(1-r)^{-1}x^2 + (1-r)^{-1}}, & \sqrt{r} \leq x \leq 1 \end{cases}. \quad (4)$$

An explanation on how the deformed circular family has been established is available in [5].

$r$  is called the *deforming parameter* and takes values in  $(0, 1)$ . Figure 1 demonstrates that values of  $r < 0.5$  squeeze the central branch of the map, while values of  $r > 0.5$  stretch it.



**Fig. 1.** Three examples of deformed circular maps and their invariant densities. (a): Map for  $r = 0.1$ , (b): map for  $r = 0.42$ , (c): map for  $r = 0.8$ , (d): density for  $r = 0.1$ , (e): density for  $r = 0.42$ , (f): density for  $r = 0.8$ .

It can be confirmed that the function  $f(x)$ , given by

$$f(x) := \begin{cases} -2(1-r)x, & -1 \leq x \leq 0 \\ 2rx, & 0 < x \leq 1 \end{cases}, \quad (5)$$

satisfies the Perron-Frobenius equation. Thus,  $f(x)$  can be seen as the invariant distribution of spreading iteratively produced by deformed circular maps. Figure 1 displays the plot of  $f(x)$  for three values of  $r$ .

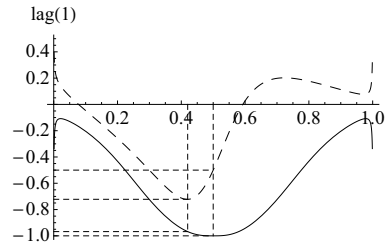
For  $r = 0.5$ , the so called *circular map* has the "V-shaped" invariant density  $f(x) = |x|$ ,  $x \in [-1, 1]$ , which results in zero mean. The mean-adjusted quadratic ACF is known for any lag (see [6]):

$$\text{Corr}(X_t^2, X_{t+s}^2) = \left(-\frac{1}{2}\right)^s, \quad s \in \mathbb{N}. \quad (6)$$

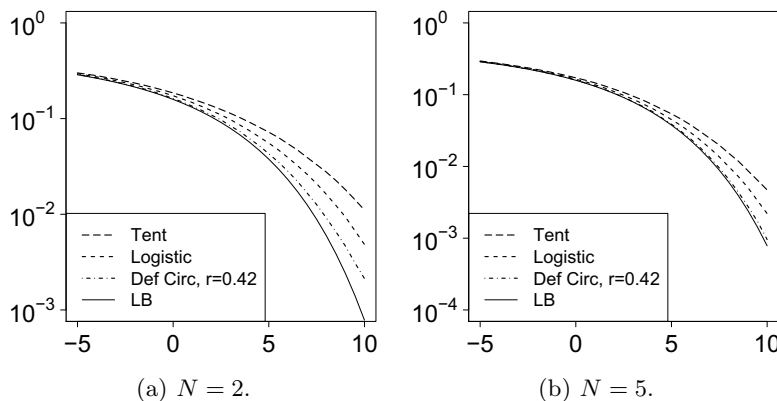
$\text{lag}(1) = -0.5$  according to (6) and therefore the circular map has smaller first lag than the tent and Bernoulli maps, which have  $\text{lag}(1) = 0.25$ .

The  $lag(1)$  value of the mean-adjusted quadratic ACF has been calculated as a function of deforming parameter  $r$ , see Figure 2. With the help of numerical minimization, it has been found that the minimal  $lag(1)$  is approximately equal to  $-0.722$  and is achieved for  $r \simeq 0.42$ . So, the optimal deformed circular map, with respect to the third optimality criterion, is the one with deforming parameter  $r \simeq 0.42$ .

**Fig. 2.** Dashed line: Plot of  $lag(1)$  of mean-adjusted quadratic ACF of deformed circular family versus its deforming parameter  $r$ . Solid line: Fréchet lower bound of  $lag(1)$  of mean-adjusted quadratic ACF for the class of spreading sequences with invariant density given by (5).



Also, the Fréchet lower bound of  $lag(1)$  has been computed, see [7]. Such a lower bound gives the minimal attainable  $lag(1)$  mean-adjusted quadratic autocorrelation among all stationary processes with invariant distribution described by (5), see Figure 2. Despite having not reached the  $lag(1)$  lower bound, spreading produced by the deformed circular maps outperforms that generated by the logistic, Bernoulli, tent and other conventionally used maps. BER simulations confirm the superiority of deformed circular family, see Figure 3.



**Fig. 3.** Simulated BER of tent, logistic, circular and deformed circular ( $r = 0.42$ ) spreading for two values of spreading factor  $N$ .

## 4 Paired Bernoulli circular spreading (PBCS)

When  $r = 0.5$ , the Fréchet lower bound of the first lag equals  $-1$ . This implies that there exists a stationary process with the "V-shaped" invariant density  $f(x) = |x|$ ,  $x \in [-1, 1]$  and  $lag(1)$  mean-adjusted autocorrelation equal to  $-1$ .

Starting from that knowledge of the distribution of spreading which fully meets the third optimality criterion, the Bernoulli circular spreading (BCS) has been defined, see [7]. Although BER simulations showed that BCS is a relatively efficient type of spreading, it has a main drawback; BCS can take only four values.

As an improvement over BCS, *paired Bernoulli circular spreading* (PBCS) has been suggested. To introduce PBCS, let  $\mathbf{X} := (X_0, X_1, \dots, X_{N-1})$  denote the  $N$ -length spreading sequence, where  $N$  is any even natural number. Also, consider the function  $f$

$$f(x) := \begin{cases} |x - k|, & x \in [k - 1, k + 1] \\ 0, & x \in \mathbb{R} \setminus [k - 1, k + 1] \end{cases}, \quad k \in \mathbb{R}, \quad (7)$$

and the random vector  $(A_1, A_3, \dots, A_{N-1})$  given by

$$A_{2i-1} := (-1)^{1-\Phi_{2i-1}}, \quad \Phi_{2i-1} \sim Ber(p), \quad i \in \left\{1, 2, \dots, \frac{N}{2}\right\}. \quad (8)$$

The even subscripted members of PBCS are defined as

$$X_{2i} \sim f, \quad i \in \left\{1, 2, \dots, \frac{N-2}{2}\right\}, \quad (9)$$

while the odd subscripted ones are

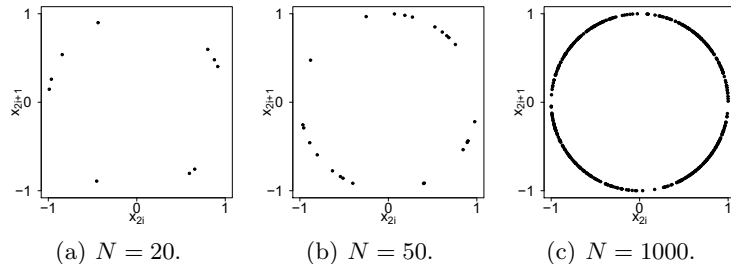
$$X_{2i-1} = k + A_{2i-1} \sqrt{1 - (X_{2i-2} - k)^2}, \quad i \in \left\{1, 2, \dots, \frac{N}{2}\right\}. \quad (10)$$

It has been shown that (7) satisfies the Perron-Frobenius equation and can thus be accepted as the invariant distribution of PBCS.

To further explain how PBCS is constructed, it could be said that the even subscripted members  $X_0, X_2, X_{N-2}$  of  $\mathbf{X}$  are sampled from (7). To obtain the odd subscripted members  $X_{2i-1}$ , a Bernoulli trial is performed to specify  $A_{2i-1}$  and then (10) is invoked in order to calculate  $X_{2i-1}$  by means of the previous value  $X_{2i-2}$ , which has been sampled from (7).

Thinking of PBCS pairwise allows for its geometrical interpretation. Each ordered pair  $(X_{2i}, X_{2i+1}), i \in \{0, 1, \dots, (N-2)/2\}$ , represents a point in a circle of centre  $(k, k)$  and radius 1, see Figure 4.

PBCS can take an infinite number of values. Geometrically speaking, any point of the unit circle with centre  $(k, k)$  could be included in PBCS. On the contrary, BCS would only allow four possible points of the circle to appear



**Fig. 4.** Simulation of PBCS from the unit circle of centre  $(0, 0)$  for spreading length  $N = 20$ ,  $N = 50$  and  $N = 1000$ .

in the spreading. So, one of the goals of PBCS has been achieved, namely to heal the weakness of BCS by allowing an infinite number of real-valued numbers in  $[k - 1, k + 1]$  to appear in the spreading.

Another question to pose is how well PBCS satisfies the third optimality criterion. The  $lag(1)$  of quadratic ACF is given by

$$Corr \left[ (X_i - k)^2, (X_{i+1} - k)^2 \right] = \begin{cases} -1, & \text{if } i \text{ is odd} \\ 0, & \text{if } i \text{ is even} \end{cases} . \quad (11)$$

So, the first lag equals  $-1$  as intended. However, this is the case only for even subscripted members of the spreading and from that point of view the third optimality criterion is not fully met.

It is interesting to point out that the invariant distribution of PBCS for  $k = 0$  coincides with the "V-shaped" density of circular map, which has deforming parameter  $r = 0.5$ . Recall that the Fréchet lower bound for the first lag has been calculated to be  $-1$  when  $r = 0.5$ . So, there has been found a stationary process whose odd subscripted members attain the Fréchet lower bound of  $lag(1)$  mean-adjusted quadratic autocorrelation.

More importantly, it has been proven that the bit energy of PBCS is constant and equal to the spreading length  $N$ . PBCS therefore meets the first optimality criterion. It has also been proven that any stationary process which meet the first optimality criterion has BER equal to the BER lower bound given in [4]. BER simulations have been run too and they verified the theory. So, one could safely state the conclusion that PBCS is optimal without leaving room for further reduction of BER.

## 5 Single-user non-coherent CSK: the model

So far optimization has focused on the transmitter and more specifically on the choice of spreading. In the coherent version of single-user CSK, there is not much to be done at the receiver end. The traditional engineering tool for decoding is the correlation decoder, which coincides with the likelihood decoder in the coherent case.

However, one may assume that the spreading  $\mathbf{X}$  is not known at the receiver, as opposed to the coherent system. Then not only  $\mathbf{T} = b(\mathbf{X} - \mu_x \mathbf{1})$ , but also  $\mathbf{X}$  has to be transmitted.

Both  $N$ -length signals  $T$  and  $X$  are corrupted by noise when passing through the channel. Eventually, two vectors  $\mathbf{R}$  and  $\mathbf{Y}$  received:

$$\mathbf{R} := b(\mathbf{X} - \mu_x \mathbf{1}) + \boldsymbol{\epsilon}, \quad (12)$$

$$\mathbf{Y} := \mathbf{X} - \mu_x \mathbf{1} + \boldsymbol{\eta}, \quad (13)$$

where it is assumed that  $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$  and  $\boldsymbol{\eta} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$ . Since  $\mathbf{T}$  and  $\mathbf{X}$  pass through the same channel, it is plausible to think of  $\boldsymbol{\epsilon}$  and  $\boldsymbol{\eta}$  having common error variance  $\sigma^2$ .  $\boldsymbol{\epsilon}$  and  $\boldsymbol{\eta}$  are also assumed to be independent of each other.

## 6 Monte Carlo Maximum Likelihood Decoding

For the non-coherent system described by (12) and (13), the correlation decoder  $C(\mathbf{y}, \mathbf{r})$  computes as the inner product  $C(\mathbf{y}, \mathbf{r}) := \langle \mathbf{y}, \mathbf{r} \rangle$ . The decoding rule based on the correlation decoder is

$$C(\mathbf{y}, \mathbf{r}) \geq 0 \Rightarrow \hat{b} = +1, \quad C(\mathbf{y}, \mathbf{r}) < 0 \Rightarrow \hat{b} = -1. \quad (14)$$

Motivated by [3], there was the thought that maximum likelihood (ML) decoding might be a more efficient alternative.

Firstly, the likelihood function  $\ell(b, \sigma^2 | \mathbf{y}, \mathbf{r})$  has been expressed as

$$\ell(b, \sigma^2 | \mathbf{y}, \mathbf{r}) = (2\pi\sigma^2)^{-N} \int \cdots \int_{D(f_{\mathbf{x}})} \exp \left[ -\frac{1}{2\sigma^2} A_{\mathbf{x}, \mathbf{y}, \mathbf{r}}(b) \right] f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}, \quad (15)$$

where

$$A_{\mathbf{x}, \mathbf{y}, \mathbf{r}}(b) := \|\mathbf{y} - (\mathbf{x} - \mu_x \mathbf{1})\|^2 + \|\mathbf{r} - b(\mathbf{x} - \mu_x \mathbf{1})\|^2. \quad (16)$$

Although the parameter of interest is  $b$ , the maximization of likelihood  $\ell(b, \sigma^2 | \mathbf{y}, \mathbf{r})$  requires the estimation of one more parameter, namely of channel noise  $\sigma^2$ . Notice that  $\ell(b, \sigma^2 | \mathbf{y}, \mathbf{r})$  is a function of the discrete variable  $b$  and of the continuous one  $\sigma^2$ .

The ML decoder  $L(\mathbf{y}, \mathbf{r})$  is defined as

$$L(\mathbf{y}, \mathbf{r}) = \max_{\sigma^2} [\ell(+1, \sigma^2 | \mathbf{y}, \mathbf{r})] - \max_{\sigma^2} [\ell(-1, \sigma^2 | \mathbf{y}, \mathbf{r})]. \quad (17)$$

and likelihood estimation imposes the ML decoding rule

$$L(\mathbf{y}, \mathbf{r}) \geq 0 \Rightarrow \hat{b} = +1, \quad L(\mathbf{y}, \mathbf{r}) < 0 \Rightarrow \hat{b} = -1. \quad (18)$$

Direct maximization of likelihood (15) is probably intractable and therefore (17) is unlikely to be calculated. A computational solution is nevertheless realizable once (15) is rewritten as

$$\ell(b, \sigma^2 | \mathbf{y}, \mathbf{r}) = E_{\mathbf{X}} \left\{ (2\pi\sigma^2)^{-N} \exp \left[ -\frac{1}{2\sigma^2} A_{\mathbf{X}, \mathbf{y}, \mathbf{r}}(b) \right] \right\}. \quad (19)$$

It was the expectation appearing in (19) which gave the idea of a Monte Carlo maximum likelihood (MCML) approach, see [2] and [1]. For the estimation of the expectation one would have to sample  $m$  spreading sequences  $\mathbf{x}_i$ ,  $i \in \{1, 2, \dots, m\}$ , each of length  $N$ . Then the Monte Carlo likelihood  $\hat{\ell}(b, \sigma^2 | \mathbf{y}, \mathbf{r})$  calculates as

$$\hat{\ell}(b, \sigma^2 | \mathbf{y}, \mathbf{r}) = \frac{1}{m} \sum_{i=1}^m \left\{ (2\pi\sigma^2)^{-N} \exp \left[ -\frac{1}{2\sigma^2} A_{\mathbf{x}_i, \mathbf{y}, \mathbf{r}}(b) \right] \right\}. \quad (20)$$

BER simulations suggest that MCML decoder  $\hat{L}(\mathbf{y}, \mathbf{r})$  outperforms correlation decoder  $C(\mathbf{y}, \mathbf{r})$ .

MCML decoding allows for exact results to be obtained. It is hoped that the idea could be implemented in more complex communication settings.

## References

1. Gelfand, A.E. and Carlin, B.P. Maximum-likelihood estimation for constrained or missing data models. *The Canadian Journal of Statistics*, 21(3):303–311, September 1993.
2. Geyer, C.J. and Thompson, E.A. Constrained monte carlo maximum likelihood for dependent data. *Journal of the Royal Statistical Society*, 54(3):657–699, 1992.
3. Hasler, M. and Schimming, T. Chaos communication over noisy channels. *International Journal of Bifurcation and Chaos*, 10(4):719–735, April 2000.
4. Lawrance, A.J. and Ohama, G. Exact calculation of bit error rates in communication systems with chaotic modulation. *IEEE Transactions on Circuits and Systems*, 50(11):1391–1400, November 2003.
5. Lawrance, A.J. and Papamarkou, T. Higher order dependency of chaotic maps. In *2006 International Symposium on Nonlinear Theory and its Applications (NOLTA)*, pages 695–698, Bologna, Italy, September 2006. Research Society of Nonlinear Theory and its Applications, IEICE.
6. Papamarkou, T. Statistical dependence of chaotic maps. Master’s thesis, University of Warwick, Department of Statistics, September 2005.
7. Papamarkou, T. and Lawrance, A.J. Optimal spreading sequences for chaos-based communication systems. In *2007 International Symposium on Nonlinear Theory and its Applications (NOLTA)*, pages 208–211, Vancouver, Canada, September 2007. Research Society of Nonlinear Theory and its Applications, IEICE.
8. Yao, J. Optimal chaos shift keying communications with correlation decoding. *Proceedings of the 2004 International Symposium on Circuits and Systems, ISCAS*, 4:IV – 593–596, May 2004.