MAXIMUM LENGTH TERNARY SIGNAL DESIGN BASED ON NYQUIST POINT MAPPING

M.F.L. Foo*, A.H. Tan** and K.P. Basu***

* Faculty of Engineering, Multimedia University, Cyberjaya, Malaysia; Email: mathias@mmu.edu.my
** Faculty of Engineering, Multimedia University, Cyberjaya, Malaysia; Email: htai@mmu.edu.my
*** Faculty of Engineering, Multimedia University, Cyberjaya, Malaysia; Email: kartik.basu@mmu.edu.my

Abstract: The design of ternary pseudo random maximum length signals suitable for system identification under noisy conditions is considered, where the power content in the specified harmonics should be high. Signal levels conversion for mapping a maximum length sequence in the Galois field into three signal levels is proposed such that the primitive version of the signal has only one nonzero harmonic at the Nyquist frequency. If the primitive signal is subsequently used in the design of a maximum length signal of a larger period, this will result in the latter having high power concentrated at the nonzero harmonics. Copyright © 2006 IFAC

Keywords: Bilinear systems, frequency spectrum, identification, input signals, model approximation, pseudo random sequences

1. INTRODUCTION

Perturbation signal design plays an important role in system identification. Signals with well designed amplitude and harmonic properties can be used in various applications to identify the underlying linear dynamics (Schoukens, et al., 2001) of a system with the effects of nonlinear distortion removed, the related linear dynamics of a system taking into account the bias caused by nonlinear distortion, as well as the identification of the nonlinearities themselves (Godfrey, et al., 2005).

Perturbation signals can be classified into computer-optimised and pseudo random signals. Computer-optimised signals are designed using a two-stage process, i.e. by specifying the required spectrum followed by optimising an arbitrary realisation of the spectrum, which takes any amplitude constraints into account (Schoukens, et al., 1993). On the other hand, pseudo random signals have fixed power spectra and their design considerations focus primarily on the signal length and the number of signal levels (Godfrey, et al., 1999). The largest class of pseudo random signals belongs to those based on maximum length sequences which exist for Galois field, GF(q), where q is a prime or a power of a prime. Such signals can be easily generated in hardware using shift registers (Tan and Godfrey, 2002) or in software using the freely available package GALOIS (Barker, 2001).

This paper focuses on the design of ternary pseudo random maximum length (PRML) signals. Such signals can be generated with odd q ≥ 3, provided the field elements e are mapped into three signal levels. It should be noted that ternary signals are particularly useful when dealing with systems where the input transducer can only accept a small number of signal levels, for example, in a hot-dip galvanising process as described by Barker and Godfrey (1999).
Three methods have been described in the literature for obtaining the signal levels conversion. The first is an analytical method, which has the advantage of giving symbolic results of great generality, as described by Barker, et al. (2004a). The second is through direct search as detailed in Barker, et al. (2004b), which allows accurate results to be obtained. The third method is a hybrid method, where the design algorithm consists of two steps. In the first step, the program multilev_new (Tan and Godfrey, 2004) is used to generate a computer-optimised primitive PRML signal. In the second step, a PRML signal of a much greater period is generated based on the signal levels conversion defined through the first step. Such a method enables the signals to be generated for large \( q \), where the first two methods become computationally inefficient (Tan, et al., 2005).

All the above three methods aim at producing signals which are inverse-repeat, having the second half period the negative of the first half, as defined in (1).

\[
u(i) = -\nu\left(i + \frac{N}{2}\right), \quad 0 \leq i \leq \frac{N}{2} \tag{1}\]

where \( i \) is the time index and \( N \) is the period.

Inverse-repeat signals contain only odd harmonic components. Dunn and Hawksford (1993) reported that these signals have the advantage of good immunity against the effects of nonlinear distortion. Such signals allow contributions of odd order and even order nonlinearities to be separated at the system output. Since inverse-repeat signals have zero even harmonic components, when these signals are used to perturb a system, even order nonlinear distortion in the system would result in nonzero even harmonic components at the output, hence providing information on the nonlinearities of the system. Examples of applications that made use of inverse-repeat signals can be found in Anzai, et al. (1999) and Tan and Godfrey (2002).

In addition to suppressing the even harmonics, harmonic multiples of three can also be suppressed to reduce the effects of odd order nonlinear distortion to the estimate of the underlying linear dynamics. In the three methods described above, the nonzero harmonics are specified to be ideally uniform for persistent excitation of the system under test.

In this paper, in contrast to the above methods, the objective is to design signals with a sparse spectrum but with the desirable property of high power content in the nonzero harmonics. In other words, the number of useable harmonics is sacrificed to achieve higher power content in the specified harmonics. It will be seen later in Section 3 that such signals have primitive versions that concentrate all their power at the Nyquist frequency alone. Hence, the mapping will be referred to here as Nyquist point mapping.

2. PSEUDO RANDOM MAXIMUM LENGTH SIGNALS

2.1 Introduction to Galois fields

Multilevel maximum length sequences exist for \( \text{GF}(q) \), where \( q \) is a prime or a power of a prime \( p \) (> 1), i.e. for \( q = 2, 3, 4, 5, 7, 8, 9, 11, 13, \ldots \) (Zierler, 1959). A GF is a collection of elements, which together satisfy the requirements of the relevant finite field theory. A PRML signal is generated from a PRML sequence by mapping the \( q \) field elements into real numbers. These signals can be generated in hardware using shift registers (Tan and Godfrey, 2002) or using the free software package GALOIS (Barker, 2001).

A pseudo random sequence, \( s_{q,n}(i) \) in \( \text{GF}(q) \) has the form of (Barker, 1993)

\[
s_{q,n}(i) = -\sum_{r=1}^{\frac{N}{2}} c_r s_{q,n}(i-r), \quad \text{all } i \tag{2}\]

where \( c_r \) are the coefficients of a primitive polynomial \( f_n(x) \) of order \( n \) in \( \text{GF}(q) \), given by

\[
f_n(x) = \sum_{r=0}^{n} c_r x^r, \quad c_0 = 1, \quad (-1)^n c_n \text{ primitive.} \tag{3}\]

A period of the pseudo random sequence is \( N = q^n - 1 \). In any one period, there are \( q^{n-1} \) occurrences of each nonzero field element and \( q^{n-1} - 1 \) occurrences of the zero field element.

2.2 Primitive pseudo random maximum length signals

A PRML sequence is primitive when \( n = 1 \) (Barker, 2004). Substituting \( n = 1 \) into (2) and (3), yields

\[
s_{q,1}(i) = -c_1 s_{q,1}(i-l), \quad \text{all } i \tag{4}\]

\[
f_1(x) = 1 + c_1 x. \tag{5}\]

Subsequently, a period of \( s_{q,1}(i) \) is \( N = q - 1 \).

The primitive PRML signal \( m(i) \) can be obtained from the primitive PRML sequence \( s_{q,1}(i) \) by the conversion of every element \( e \) of \( \text{GF}(q) \) into a signal level \( \nu(e) \) such that (Tan, et al., 2005)

\[
m(i) = \nu(s_{q,1}(i)), \quad \text{all } i \tag{6}\]

and in a period of \( m(i) \)

\[
m(i) = \nu(q^{-i}), \quad i = 1, 2, \ldots, q - 1 \tag{7}\]
where \( g \) is a primitive element of the field, having the important property that its powers \( g^0, g^1, \ldots, g^{q-2} \) generate all the nonzero elements of the field.

In a similar way, a PRML signal \( u(i) \) with period \( N = q^n - 1 \) can be generated from \( s_{q,n}(i) \) using the same signal levels conversion as for the primitive PRML signal.

\[
u(i) = v(s_{q,n}(i)) \text{, all } i \quad (8)
\]

The harmonic content of the discrete signal \( u(i) \) is defined by the magnitude of its discrete Fourier transform (DFT) given by

\[
|U(k)| = \left| \sum_{i=0}^{N-1} u(i) \exp \left( -\frac{2\pi j i k}{N} \right) \right| 
\]

\[
(9)
\]

where \( k = 0, 1, \ldots, N - 1 \).

Provided \( q \) is odd and \( v(0) = 0 \), \( |U(k)| \) is related to the magnitude of the DFT of \( m(i) \), denoted by \( |M(k)| \), through

\[
|U(k)| = |M(0)| q^{-1}, \quad k \text{ a multiple of } q^n - 1; \quad (10A)
\]

\[
|U(k)| = |M(0)| q^{(n-2)/2}, \quad k \text{ a multiple of } q - 1 \text{ but not of } q^n - 1; \quad (10B)
\]

\[
|U(k)| = |M(k)| q^{(n-1)/2}, \quad k \text{ not a multiple of } q - 1. \quad (10C)
\]

This allows a PRML signal to be designed through the design of its primitive version, which has a shorter period (Barker, 2004).

3. SEQUENCE TO SIGNAL CONVERSIONS

3.1 Mapping for uniform odd harmonics with even harmonics suppressed

A PRML sequence from any higher field \( q \) can be mapped into a PRML signal with any lower number of signal levels \( \leq q \) (Barker, et al., 2004a). To preserve uniformity of odd harmonic components and suppress even harmonic components, the signal must fulfil (11) and (12),

\[
m(i) + m \left(i + \frac{q-1}{2}\right) = 0 \quad (11)
\]

\[
\varphi_{\text{autocorr}}(i) = 0, i \neq r \left\lfloor \frac{q-1}{2} \right\rfloor \quad (12)
\]

where \( \varphi_{\text{autocorr}}(i) \) is the periodic autocorrelation, and \( r \) is an integer.

Details are given in Barker, et al. (2004a) on how such a mapping can be obtained. For example, the signal levels conversions to map a GF(5) sequence into a ternary signal are shown in Table 1, where \( a \) represents a real number.

| Table 1: Mapping a 3-level signal from GF(5) |
|---------------|---------------|---------------|---------------|
| \( m(1) \)   | \( m(2) \)   | \( m(3) \)   | \( m(4) \)   |
| Option 1     | \( a \)      | \( 0 \)      | \( -a \)     | \( 0 \)      |
| Option 2     | \( a \)      | \( a \)      | \( -a \)     | \( -a \)     |

3.2 Nyquist point mapping

For signals with the same amplitude distribution, it is possible to increase the power in the nonzero harmonics by making the spectrum sparser. A mapping is proposed in this paper where the primitive PRML signal \( m(i) \) has all the power concentrated on the Nyquist frequency. This can be achieved with

\[
m(i) = a, i \text{ odd}; \quad (13A)
\]

\[
m(i) = -a, i \text{ even.} \quad (13B)
\]

Arbitrarily setting \( a = 1 \) gives

\[
|M(k)| = q^{-1}, k = \left\lfloor \frac{q-1}{2} \right\rfloor + r(q-1); \quad (14A)
\]

\[
|M(k)| = 0, k \text{ otherwise} \quad (14B)
\]

where \( r = 0, 1, 2, \ldots \).

The DFT of \( |U(k)| \) can then be obtained using (10) and (14). As an example, for \( q = 5 \), using \( m(i) = [+1, +1, -1, -1] \) according to Option 2 in Table 1 results in the magnitude of the DFT as shown in Figure 1 (top). However, with the Nyquist point mapping, \( m(i) = [+1, -1, +1, -1] \) and this gives the DFT magnitude as plotted in Figure 1 (bottom). Comparing these, it can be seen that the former has two nonzero harmonics in a period of \( m(i) \), while the latter has only one nonzero harmonic. However, the power in the nonzero harmonic in the latter is twice larger than that in the former.

To illustrate the relationship between \( |U(k)| \) and \( |M(k)| \), PRML signals with \( n = 3 \) and \( N = 124 \) are generated from the primitive signals. The resulting PRML signals are referred to here as Signal 1 (generated using the mapping in Table 1) and Signal 2 (generated using the Nyquist point mapping). Their DFT magnitudes are shown in Figure 2.
4. APPLICATION EXAMPLE: ESTIMATION OF RELATED LINEAR DYNAMICS OF FIRST ORDER BILINEAR SYSTEMS

4.1 Introduction to first order bilinear systems and simulation settings

A first order bilinear system has the form of

$$\dot{z} = -\alpha z + u + \rho z u, \ y = \beta z$$

(15)

where $u$ is the input; $y$ is the output; $z$ is the system state; and $\alpha$, $\beta$ and $\rho$ are constants. Such a system is nonlinear due to the multiplicative term involving the input and system state. Examples in the industry are gas-fired furnaces (Dunoyer, et al., 1997) and polymerisation reactors (Xu, et al., 1994). The related linear dynamics can be used as a linear approximation to the system, as these contain the underlying linear dynamics plus the effects of bias due to nonlinear distortion. For first order bilinear systems, the related linear dynamics are measured in terms of the related time constant, $T_C$, which can be calculated theoretically using (Tan, 2005) as

$$T_c = -\frac{T}{\ln\left(e^{-T\alpha} + e^{-T\beta} + e^{-T\rho} - 2\right)}$$

(16)

where $R$ is the number of elements in GF($q$) mapped into signal levels $\pm V$ and $T$ is the sampling interval.

In this simulation, $\alpha$ and $\beta$ are set to 1 while $\rho$ is varied, hence varying the amount of nonlinear distortion in the system. A Gaussian noise with zero mean and giving a signal-to-noise ratio (SNR) of 10dB is added to the system input and output, respectively, in two separate experiments. Signal 1 and Signal 2 are then used to perturb the noisy bilinear system to evaluate the closeness of the estimated related time constant to its theoretical value. For the signals used, two field elements are converted into signal level $+1$, and two field elements into signal level $-1$. Using $q = 5$, $R = 4$ and $V = 1$, (16) can be simplified and rewritten as

$$T_c = -\frac{T}{\ln\left(e^{-T\alpha} + e^{-T\beta} + e^{-T\rho} - 2\right)}$$

(17)

4.2 Identification results

The output DFT magnitudes using Signal 1 and Signal 2 are illustrated in Figure 3 for the bilinear system with $\rho = 0.5$. It can be seen from Figure 3 that the estimated frequency response using Signal 1 is less smooth due to the effects of noise.
Fig. 3. Output DFT magnitudes of the bilinear system with $\rho = 0.5$. For the sake of clarity, only the first half period is plotted. Top: Signal 1; Bottom: Signal 2.

The Estimator for Linear Systems (ELiS), available from the Frequency Domain System Identification Toolbox (Kollár, 1994) in MATLAB is used to obtain the estimates of the related time constants. The simulation is run 100 times. The averaged results are shown in Figure 4.

From Figure 4, it is observed that in general, for both input and output noise, perturbation using Signal 2 gives the estimated related time constants closer to the theoretical results for different values of $\rho$ compared with those using Signal 1. This is due to the higher power content in the nonzero harmonics of Signal 2, which provides the signal with greater robustness towards noise. It should be noted that both Signal 1 and Signal 2 have the same value of Performance Index for Perturbation Signals (PIPS) (Godfrey, et al., 1999) defined by

$$\text{PIPS} = \frac{200}{N} \left( \frac{\mu_{\text{max}} - \mu_{\text{min}}}{\sum_{k=1}^{N-1} |U(k)|^2} \right)^{1/2} \% \quad (18)$$

which is 89.8%. In (18), $u_{\text{max}}$ and $u_{\text{min}}$ are the maximum and minimum values, respectively, of $u(i)$.

To confirm the results obtained using Gaussian noise, the simulation is repeated using random noise of zero mean and giving an SNR of 10dB. The averaged results are shown in Figure 5.

In general, Signal 2 gives slightly larger values of the standard deviation for the estimated related time constants, for both Gaussian noise and random noise. This could be due to the fact that estimation using Signal 2 relies on fewer number of harmonics, and changes in any estimated harmonic have a larger overall effect on the results obtained.
5. CONCLUSION

A method to convert a PRML sequence into a ternary PRML signal has been proposed. The primitive signal is mapped such that it alternates between the maximum and minimum signal levels, thus resulting in power only at the Nyquist frequency. When the primitive signal is applied in the design of a PRML signal with a greater period, the latter has a large number of harmonics suppressed, but with high power in the nonzero harmonics.

The effectiveness of the proposed mapping method was illustrated on first order bilinear systems where the signal generated using the Nyquist point design is more robust against the effects of noise compared with a design with uniform odd harmonics and suppressed even harmonics. The proposed mapping has potential applications in autotuning, where normally a small number of harmonics is sufficient for the identification.

ACKNOWLEDGEMENTS

Financial support from the Ministry of Science, Technology and Innovation, Malaysia under Intensification of Research in Priority Areas (IRPA) Grant No: 03-99-01-0058-EA054 is gratefully acknowledged.

REFERENCES


