The Consultative Committee for Space Data System
Space Link Services Area
Space Link Coding & Synchronization Working Group

SLS-CS 18-04

Randomizer for High Data Rates

H. Garon[1], V. Sank[1]
[1]ASRC (AS&D)
Work performed for NASA Goddard Space Flight Center, Greenbelt, MD 20771

CCSDS SLS-RFM & SLS-C&S WG meeting
11-12 April 2018

Introduction

NASA, as well as a number of other space agencies, now recognize that the current recommended CCSDS randomizer [1,2] used for telemetry (TM) is too short. When multiple applications of the PN8 Maximal Length Sequence (MLS) are required in order to fully cover a channel access data unit (CADU), spectral problems in the form of elevated discrete spurious (spurs) appear [3]. At the very least, the appearance of spurs portends the violation of international spectrum regulations. More importantly, the probability of false decoder lock increases dramatically. Originally, the randomizer was called a bit transition generator (BTG) precisely because it was thought that its primary objective was to ensure sufficient bit transitions to allow the bit/symbol synchronizer to lock and remained locked. We, NASA, have demonstrated that the original BTG concept is a limited view of the real value of the randomizer sequence and that the randomizer also aids in signal acquisition as well as minimizing the potential for false decoder lock.

In 1994, a PN8 MLS was selected based on avoiding with a sequence that also appeared as a Reed-Solomon RS(255,223) codeword. Apparently, in developing a recommendation for a Telecommand (TC) uplink randomizer, it was discovered that the pre-1994 recommended PN8 randomization sequence was a legitimate member of the set of RS(255,223) codewords. Present day channel access data unit (CADU) construction should, by default, preclude that possibility. Nonetheless, we'll still retain that negative property as a characteristic of any new candidate MLS we'll propose. Since there is no interaction between Telemetry (TM) and TC links, there is no need to modify the current CCSDS randomization recommendation when applied to TC.

CCSDS limits uncoded transfer frames and generally codeblocks to a maximum length of 2048 bytes (16384 bits) under most CODEC circumstances. There are exceptions. A rate 1/6 turbo code for the longest transfer frame requires 53544 bits. Similarly, a rate 1/2 LDPC will require a codeblock constructed using 32768 bits for the longest transfer frame. However, we found it reasonable to restrict our analysis here to review PN MLS performance with a maximum frame length of 16384 bits. It will become clear later that limiting our performance review to 16384 bits will neither restrict our candidates in any manner nor have any substantive effect upon our final recommendation.

To be published in the proceedings of the Consultative Committee for Space Data Systems (CCSDS)
Spring Technical Meetings, Gaithersburg, MD, April 9-13, 2018.

Procedure

In developing a new randomization recommendation, we employed the following guidelines:

- The sequence must be a Maximal Length Sequence (MLS) under Galois Field, GF(2). This implies that any recommendation would meet and maintain fundamental algebraic properties critical to maintaining pseudo-noise performance with respect to auto- and cross-correlation.
- The MLS must be constructed using a minimal number of feedback register terms. Any terms beyond minimal translates to additional hardware and/or software required for implementation.
- The longest CADU blocklength that will be considered in this analysis will be 2^{14} (16384) bits. The shortest blocklength will be 2^8-1 (255) bits.
- No constraints are placed on the shift register initialization, regardless of order. For our immediate purposes here, we assume all registers of the generator will be set (initialized to '1').

However, we do recognize the potential for further optimization by carefully selecting the MLS segment that will be used. PN sequences with length greater than our maximum blocklength of 214 will have that advantage over shorter length sequences. The portion selection would be accomplished by initializing the registers to start the sequence at the desired segment. Using code scripts written under Mathworks MATLAB®, we first generated all primitive polynomials under GF(2) having degrees (M) 8 through 23, inclusively. Since the definition of a primitive polynomial requires irreducibility, all primitive polynomials share the exclusive feature of also being maximal length sequence generators, i.e., for a primitive polynomial of degree M, the generator sequence will not repeat itself for 2M-1 shifts. We further restrict the candidates to a single allowable feedback term. This implies that all of our candidates will appear in the form:

$$PN(x) = 1 + x^p + x^M,$$

where M is again the degree of the polynomial and p is the selected feedback tap such that $1 \leq p < M$. Restricting the generator to a single feedback term reduces the number of viable candidates to twenty-eight (28). The twenty-eight candidates are actually comprised of fourteen pairs, with each member of the pair related to each other by the inverse operator $PN^{-1}(x) = x^M PN(x^{-1})$.

The twenty-eight candidates then underwent scrutiny for their auto-correlation performance. In a strict sense, the auto-correlation is a demonstration of a sequence's characteristic ability to appear pseudo-random over the full length of the sequence. Since the only sequences we considered are primitive, they are, by virtue of being primitive, all maximal length sequences (MLS). In a truly random sequence, the only non-zero correlation coefficient in the auto-correlation function will have zero lag. In a pseudo-random sequence, the zero lag coefficient dominates with the remaining coefficients being finite but small. A respectable pseudo-random sequence will have the zero lag coefficient close to one and all the remaining coefficients in the neighborhood of two orders of magnitude down.

We immediately excluded all sequences with polynomial degree $M < 14$. Any sequence which falls short of covering our maximum considered frame length (2^{14}) required multiple applications of that sequence. As an example, consider the current recommended MLS,

$$PN_C(x) = 1 + x^3 + x^5 + x^7 + x^8.$$

The autocorrelation over lags l for $-255 < l < +255$ appears in figure 1a. The vertical axis (correlation coefficient $C_{aa}(l)$) is expanded in figure 1b to show the details of the correlation coefficient for non-zero

lag. Peaks on the order of +/- 0.02 indicate the pseudo-noise is reasonable, i.e., the non-zero lag correlation coefficients are 13 dB down from the zero-lag component. When the $PN_c(x)$ sequence is applied repeatedly over a frame length of 2^{14}, the autocorrelation over all possible lags appears in figure 2 where now the autocorrelation is filled with spikes associated with the repeat cycle. Similarly, increasing the degree of the MLS to 10, results in fewer spikes (fewer repetitions) but still the behavior is unacceptable. Figure 4 shows the autocorrelation for the first three-term MLS with degree ≥ 14,

$$PN(x) = 1 + x^8 + x^{15}.$$

Here the skirts of the autocorrelation are on the order of 20 dB down from the zero-lag component.

In order to support our discussion of the remaining candidate sequences, we'll need another performance metric for comparing sequences, and one that will illustrate the autocorrelation performance as a function of varying frame lengths. Towards that end, the metric we propose,

$$\sum_{l=-L}^{+L}[C_{aa}(l) \cdot C_{aa}(l)]/L^2,$$

is constructed for each and every L such that 256 ≤ L ≤ 16384. $C_{aa}(l)$ is the correlation coefficient and its square $[C_{aa}(l) \cdot C_{aa}(l)]$ is often referred to as the coefficient of determination. An increasing coefficient of determination implies a decreasing variability or, in our application, poorer performance in the skirts of the auto-correlation as a pseudo-random generator. Using this metric, the performance of a particular MLS for all possible frame lengths L for 256 ≤ L ≤ 16384 is made dramatically evident. Consider figure 5 where four candidate sequences are compared with each other. One of those sequences, in particular

$$PN(x) = 1 + x^7 + x^{10},$$

initially demonstrates a diminishing metric and then, once the frame length exceeds the sequence length ($L > 2^{10}$), it begins increasing dramatically compared to the other sequences. All the other candidates having greater degree exhibit much more promising behavior and begin to approach an asymptote with increasing L.

Discussion of Candidate Sequences

While MLS of degree > 14 (our maximum frame length) exhibits asymptotic behavior, increasing degree also results in an increase in the asymptote. In other words, MLS with degree close to the maximum frame size minimize the metric asymptote. Our goal is to select the MLS that would simultaneously minimize the asymptote and still provide optimized performance for smaller frame lengths. Figure 6a illustrates that tradeoff in MLS performance. Three MLS of degree 15 are compared with one at degree 17. All four monotonically approach an asymptote once beyond a frame length of 3500. As explicitly shown in figure 6b, however, each sequence distinguishes itself prior to that frame length of 3500 by its own unique behavior. Two of the sequences have comparatively large metrics for frame lengths L for 255 < L < 1250. This suggests that they would perform poorer than the other two for selecting L over this limited domain. It appears that the last MLS, $PN(x) = 1 + x^3 + x^{17}$, offers the best compromise among all the MLS plotted in the figure. Note that the metric performance of $1 + x^7 + x^{15}$ was a close second.

ESA had addressed this same problem in a series of papers and presentations between 2008 and 2009 [4-6]. The authors unambiguously established the mechanism for the spurious content and recognized as well that a longer MLS was required. The ESA proposal was to use $1 + x^{14} + x^{15}$. ESA knew that their recommendation had distinct problems when applied to frame lengths much shorter than the sequence

itself. That shortcoming is seen in figures 6a and 6b when compared to the other candidate sequences. However, not too long after their MLS proposal, ESA recommended [7] that the registers could be initialized to [100101010000000]. Figure 7, taken from the European Telecommunications Standards Institute (ETSI) draft document [7], removes any ambiguity regards to interpretation of the initialization. Initializing the registers to something other than all ones just shifts where the sequence begins and does not affect the sequence length. In fact, proper choice of initialization can shift sequence portions with greater repetitive weights towards the end of the sequence rather than the beginning of the sequence. The ESA suggestion of an initialization sequence other than all ones attempts precisely that. As shown in figure 8, modifying the initialization does dramatically improve the performance of the MLS for the shorter sequences.

Figures 9, 10 and 11 show additional comparisons between various candidates using the metric. Figure 10 plots the metric for what we consider are the four finalists. The autocorrelation (figure 12) over a frame length of 2^{14} of one of the PN17 finalists may be contrasted with one of the PN15 finalists (figure 4).

Recommendation and Summary

Under the guidelines we considered here, there are multiple maximal length sequences under GF(2) which appear attractive. Although there may be mitigating reasons why another MLS sequence would be selected, one sequence in particular,

$$PN(x) = 1 + x^3 + x^{17},$$

possesses a combination of desired properties which offsets it from the others. The autocorrelation for this MLS is shown in figure 10 and is quite similar to the autocorrelation shown in figure 4 for $1 + x^8 + x^{15}$. The polynomial is very simple in construction, employs only three terms (one feedback register), and already displays close to optimum auto-correlation properties with all shift registers initialized to one. Coupled with the prospect that this sequence length (2^{17}-1) can anticipate much greater frame lengths beyond our restrictions here, these properties make the sequence attractive to promulgate and implement in practice. Our new recommendation should aid in signal acquisition by supporting bit/symbol synchronization as well as minimizing decoder false lock.

References

[1] *TM Synchronization and Channel Coding*. Recommendation for Space Data System Standards, CCSDS 131.0-B-3. Blue Book. Issue 3. September 2017.

[2] *TM Synchronization and Channel Coding – Summary of Concept and Rationale.* CCSDS 230.1-G-2. Green Book. Issue 2. November 2012.

[3] O. Alvarez and G. Lesthievent, "Pseudo-Random Codes for High Data Rate Telemetry: Analysis and New Proposal", CCSDS RF & Modulation & Channel Coding Working Groups, Roma, 12 June 2006.

[4] M. Baldi, G. P. Calzolari, F. Chiaraluce, and R. Garello, "Randomizer for High Data Rates: Some Proposals against the Problem of Spectral Spurious", CCSDS Coding & Synchronization Working Group, Fall Meeting, Berlin, 13 October 2008.

[5] R. Garello, M. Baldi, F. Chiaraluce, "Randomizer for High Data Rates", Politecnico di Torino, European Space Agency Contract Report, ESOC Contract No. 20959/07/D/MRP, March 2009.

[6] R. Garello, M. Baldi, G. P. Calzolari (ESA/ESOC), "Summary of High Data Rate Randomizer Investigations", CCSDS SLS-Coding & Synchronization Working Group, Spring Meeting, Colorado Springs, 21 April 2009.

[7] "Draft ETSI EN 302 307 V1.3.1 (2012-11)", European Telecommunications Standards Institute (ETSI), Sophia Antipolis Cedex, 2012.

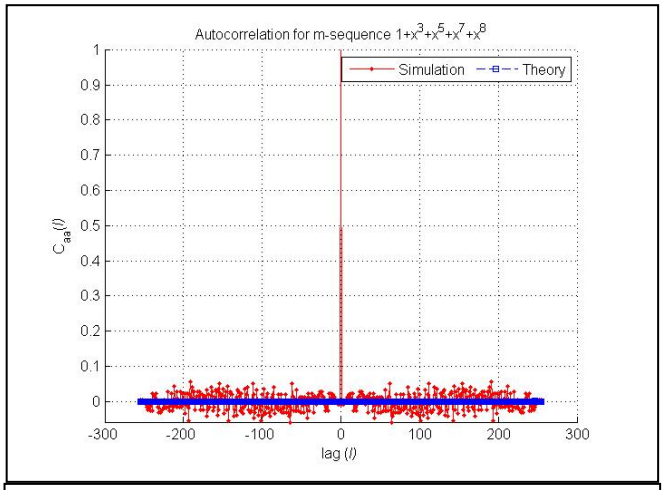

Figure 1a. Autocorrelation for current recommended PN8 MLS randomizers.

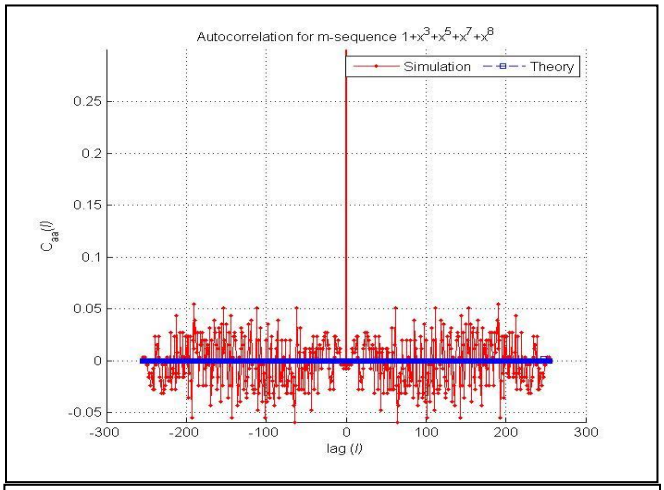

Figure 1b. Expanded correlation axis to show details of autocorrelation (current PN8 MLS).

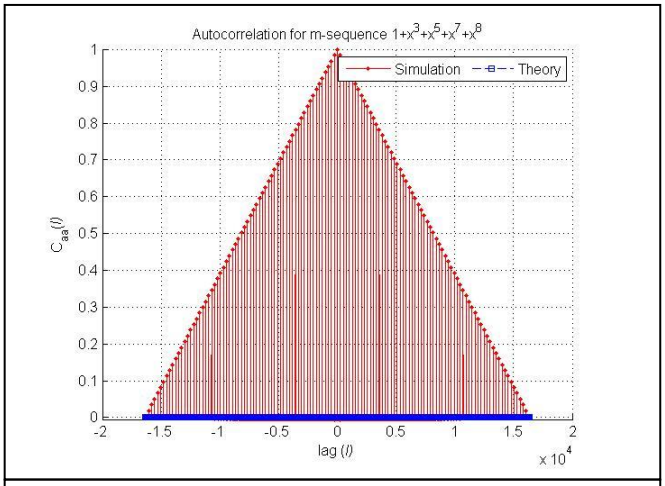

Figure 2. Autocorrelation of repeated application of PN8 MLS across frame length of 2^{14} points.

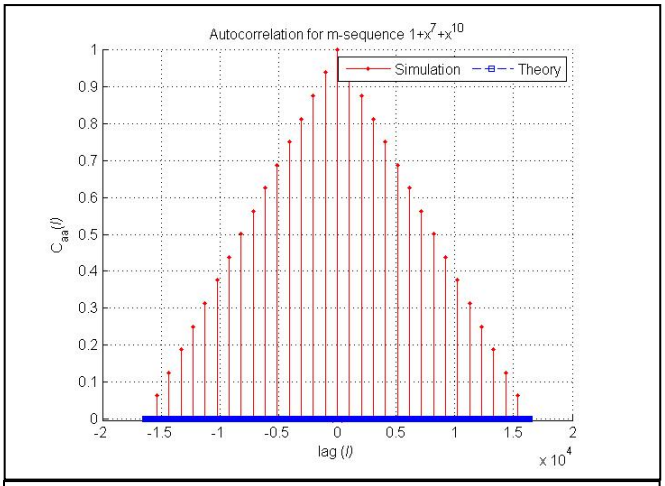

Figure 3. Autocorrelation of repeated application of PN10 MLS across frame length of 2^{14} points.

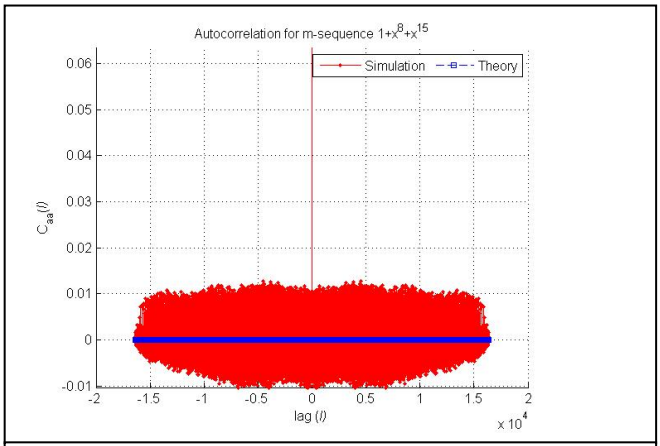

Figure 4. Autocorrelation of selected PN15 MLS: $1+x^8+x^{15}$ across frame length of 2^{14} points.

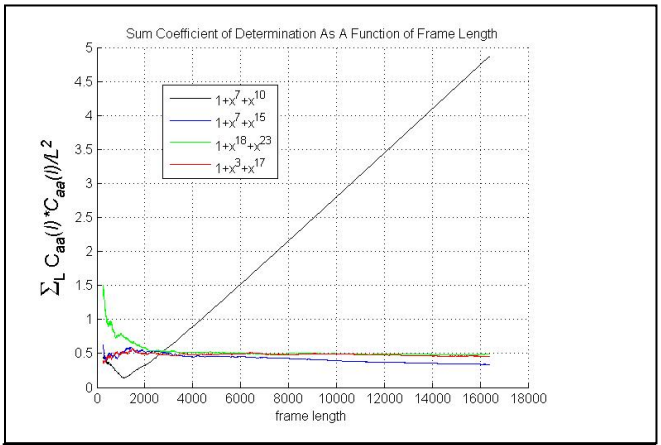

Figure 5. Metric of choice: Sum coefficient of determination as a function of frame length.

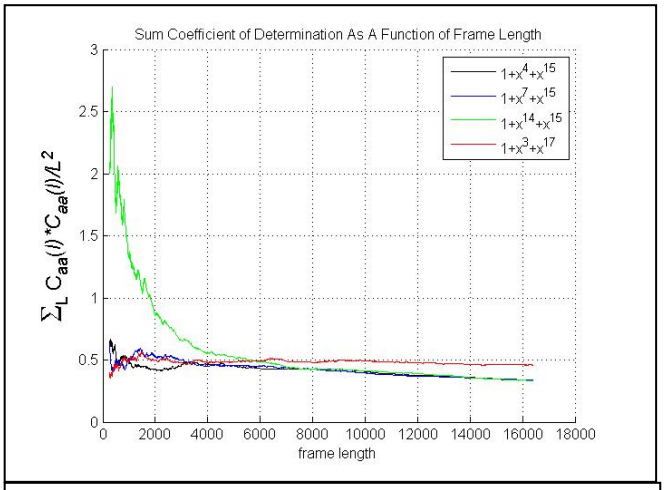

Figure 6a. Sum coefficient of determination as a function of frame length.

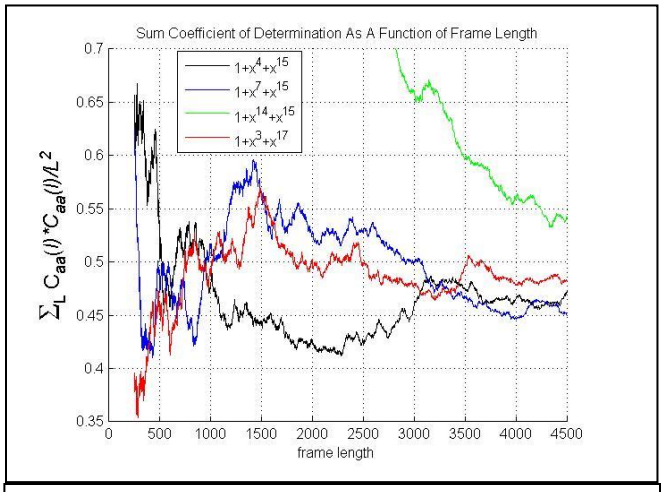

Figure 6b. Expanded axes to show details of Sum Coefficient of Determination.

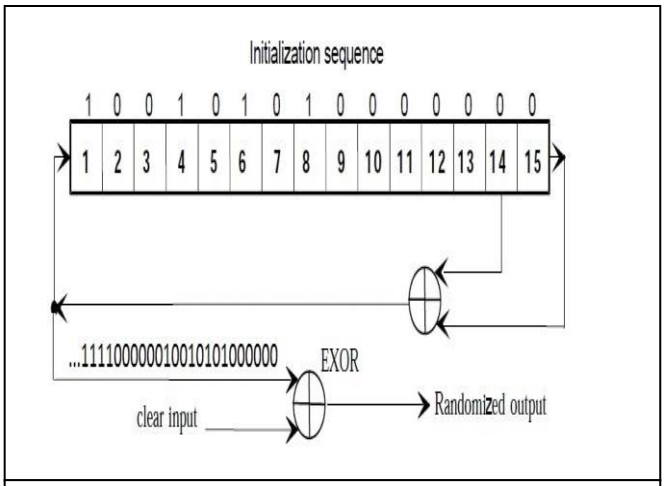

Figure 7. CNES recommendation with requisite initialization (Figure derived from ref [7].)

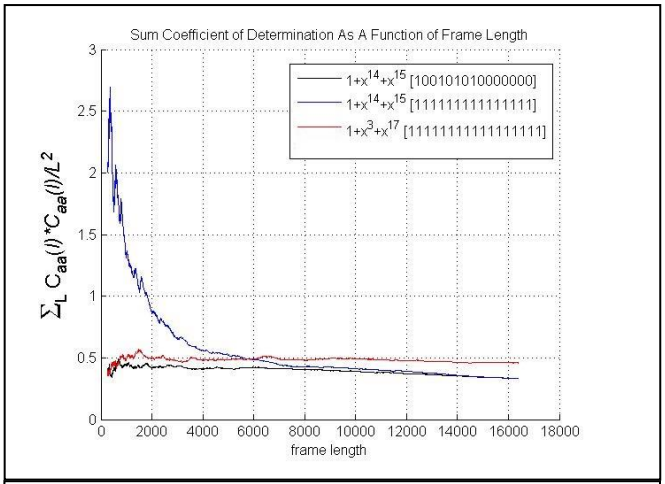

Figure 8. Sum coefficient of determination as a function of frame length (CNES recommendation).

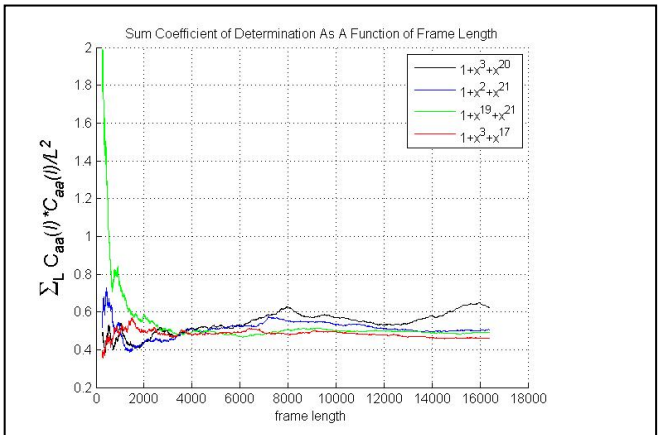

Figure 9. Sum coefficient of determination as a function of frame length

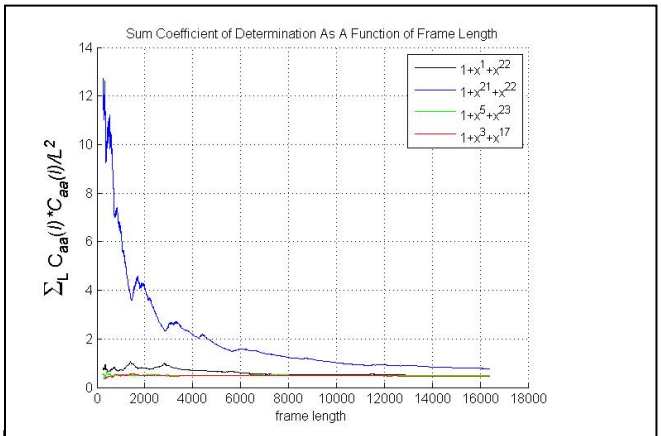

Figure 10. Sum coefficient of determination as a function of frame length.

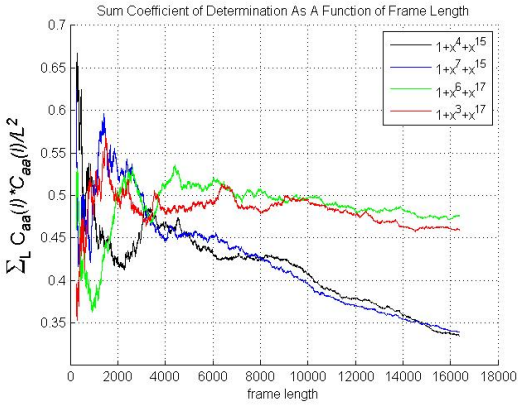

Figure 11. Four finalists not requiring special initialization.

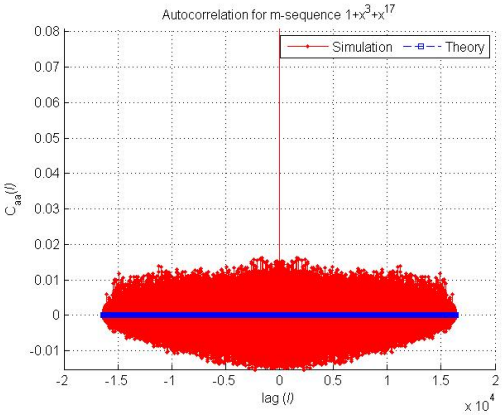

Figure 12. Autocorrelation of selected PN17 MLS: $1+x^3+x^{17}$ across frame length of 2^{14} points.

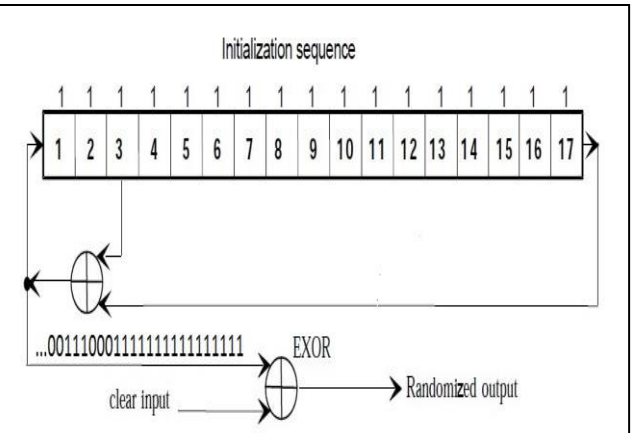

Figure 13. Recommended MLS ($1+x^3+x^{17}$) for randomization along with suggested initialization.