Binary sequences with merit factor greater than 6.34

Peter Borwein Kwok-Kwong Stephen Choi
Jonathan Jedwab
Department of Mathematics, Simon Fraser University,
Burnaby, BC, Canada V5A 1S6

23 July 2004

Abstract

The maximum known asymptotic merit factor for binary sequences has been stuck at a value of 6 since the 1980s. Several authors have suggested that this value cannot be improved. In this paper we construct an infinite family of binary sequences whose asymptotic merit factor we conjecture to be greater than 6.34. We present what we believe to be compelling evidence in support of this conjecture. The numerical experimentation that led to this construction is a significant part of the story.

Keywords aperiodic autocorrelation, asymptotic, binary sequence, merit factor

1 Preamble

We begin with a general discussion of the merit factor problem for binary sequences and outline our approach, prior to a more formal introduction in Section 2. The problem of determining the maximal merit factor for binary sequences is an old and apparently very difficult problem in combinatorial optimization — or perhaps in analytic number theory, depending on where the eventual solution lies.

The computational problem involves searching amongst sequences of length n having elements -1 or 1 to determine the maximal merit factor. The search space is therefore of size 2^n, and the merit factor of any given sequence can be computed algebraically from its elements in $O(n^2)$ operations. The combinatorial "landscape" of the search space appears to be very rugged, which is particularly challenging for stochastic search methods. On the other hand exhaustive search methods have so far been carried out only as far as length 60 [1], while a length of 100 is well beyond the range of current methods and machines. Indeed, this "... seems to be amongst the most difficult optimization problems" [2].

More than twenty years ago Turyn, as reported by Golay [3], observed that by cyclically rotating Legendre sequences we can obtain sequences of increasing length whose merit factor appears to converge to 6. A proof of this was given in 1988 by Høholdt and Jensen [4]. Other explicitly constructed families of sequences have been shown to have non-zero asymptotic merit factor but never with a value greater than 6. This is essentially where the problem has been stuck, and several authors have cast doubt on any improvement being found. For example in 1983 Golay [3] stated that:

> "[Six] is the highest merit factor obtained so far for systematically synthesized binary sequences, and the eventuality must be considered that no systematic synthesis will ever be found which will yield higher merit factors."

while in 1988 Høholdt and Jensen [4] declared:

> "We therefore make a new conjecture concerning the merit factor problem, namely, that asymptotically the maximum value of the merit factor is 6 and hence that offset Legendre sequences are optimal."

In 1999 Tony Kirilusha and Ganesh Narayanaswamy were supervised as summer students at the University of Richmond by Jim Davis. They observed [5] that, beginning with Turyn's sequences having asymptotic merit factor 6, they could append the initial part of the sequence to the end of the sequence in order to obtain sequences with merit factor clearly greater than 6. They speculated that in general this behaviour would be exhibited for a Turyn sequence of length n, by appending $\lfloor n^\alpha \rfloor$ sequence elements in this way for some fixed value of $\alpha < 1$. In fact this speculation is not correct, as we shall prove in Proposition 6.1, but even after establishing this we had the instinct that there was a real phenomenon here. It turns out that Kirilusha and Narayanaswamy's insight was a very good one but their computations did not involve sufficiently long sequences. We found it necessary to examine sequences of length in the tens of thousands in order to get a clear picture of the behaviour they discovered, and to go to lengths well in excess of one million (where the merit factor computation for a given sequence is of the order of 10^{12} operations) in order to find sufficient numerical evidence to formulate an explanatory conjecture. The calculations, while not extraordinarily hard, require careful orchestration and attention to efficiency of implementation. Certainly Maple or Mathematica will not suffice.

The key turns out to be that the number of sequence elements to be appended should be a fixed multiple $\lfloor tn \rfloor$ of the sequence length n rather than a fractional power $\lfloor n^\alpha \rfloor$. Our analysis firstly expresses the merit factor of an appending of a rotated Legendre sequence in terms of the merit factor of certain truncated rotated Legendre sequences (see Theorem 6.4). It is then sufficient to determine the merit factor of any truncation of any rotation of a Legendre sequence. After much computation, wrong guessing, and curve fitting we uncovered, at least numerically, a formula involving this merit factor (see Conjecture 7.5). Subject to the correctness of this conjecture we can explain precisely the phenomenon observed in [5] and obtain a family of binary sequences with merit factor greater than 6.34 (see Corollary 8.2). The most convincing evidence we have in support of Conjecture 7.5 is that represented in Figures 6 and 8.

We will present the results of our computational explorations in some detail in order to illustrate the discovery process that links the work of Kirilusha and Narayanaswamy [5] to Theorem 6.4, Conjecture 7.5, and Corollary 8.2.

2 The Real Introduction

A binary sequence A of length n is an n-tuple $(a_0, a_1, \ldots, a_{n-1})$ where each a_i takes the value -1 or 1. The *aperiodic autocorrelation* of the sequence A at shift u is given by

$$C_A(u) = \sum_{i=0}^{n-u-1} a_i a_{i+u} \qquad \text{for } u = 0, 1, \ldots, n-1. \tag{1}$$

The study of sequences whose aperiodic autocorrelations ($C_A(u)$) are, in some suitable sense, small is one of the classical problems of sequence design. Such sequences have been sought since the 1960s for digital communications applications such as synchronization, position determination, and pulse compression. In fact Turyn's interest in this problem [6], and in particular the Barker sequence problem [7], appears to have been one of the motivations for his celebrated paper [8] that established the systematic use of algebraic number theory in the study of difference sets (see [9] for an overview of this now standard technique, and some precursors to [8]).

The *merit factor* of a sequence A of length n is

$$F(A) = \frac{n^2}{2\sum_{u=1}^{n-1}[C_A(u)]^2}, \tag{2}$$

as introduced by Golay [10]. A sequence whose aperiodic autocorrelations are collectively small, as measured by the sum of their squares, has a large merit factor. Let $\mathcal{A}_n$ be the set of all binary sequences of length n. We define F_n to be the optimal value of the merit factor for sequences of length n:

$$F_n = \max_{A \in \mathcal{A}_n} F(A).$$

The two principal problems in the study of the merit factor are:

A. Determine F_n for a given value of n

B. Determine the asymptotic behaviour $\limsup_{n \longrightarrow \infty} F_n$.

We believe Problem B to be the more fundamental, though clearly a solution to Problem A for all values of n would solve Problem B. A particular case of Problem B is to show whether $\limsup_{n \longrightarrow \infty} F_n$ is bounded above; if so this would imply a conjecture due to Erdös [11] dating from 1962 [12, p. 127], as well as the long-standing Barker sequence conjecture (see [13] for recent progress on this conjecture). By an easy parity argument we know that $F_n \leq n^2/(n-1)$ for odd n and $F_n \leq n$ for even n; by [7] equality cannot be achieved here for odd $n > 13$, and the Barker sequence conjecture is equivalent to equality being unattainable for even $n > 4$.

With regard to Problem A, the largest known values of F_n are $F_{13} = 169/12 \simeq 14.1$ and $F_{11} = 12.1$. No $n > 13$ is known for which $F_n \geq 10$. The value of F_n was determined by Mertens for $n \leq 48$ [14], and subsequently by Mertens and Bauke for $n \leq 60$ [1], using an exhaustive branch-and-bound algorithm with an apparent running time of

approximately 1.85^n. Lower bounds on F_n for some values of $n > 60$ have been found by exploring a restricted search space, for example the set of "skew-symmetric" sequences of odd length n [15], [16]. Recent numerical investigations have shown that $F_n > 7.7$ for all n in the range $41 \leq n \leq 185$ [17].

With regard to Problem B, Mertens [14] applied curve-fitting to the value of F_n for $n \leq 48$ to suggest that $\limsup_{n\longrightarrow\infty} F_n > 9$. Golay [18] gave a heuristic argument, based on an assumption he termed "the ergodicity postulate", whose conclusion is that $\limsup_{n\longrightarrow\infty} F_n \simeq 12.3$. Golay's argument was presented in a simpler form by Bernasconi [19] although we still do not find the underlying assumption convincing. The best proven asymptotic result is that $\limsup_{n\longrightarrow\infty} F_n \geq 6$. This result is due to Høholdt and Jensen [4], based on numerical work carried out by Turyn (as reported in [3]) and developed by Golay [3]. The method of [4] is to show that there is a family of sequences of prime length n for which the merit factor tends to 6. This family is an essential ingredient in the construction presented in this paper. Although the method of [4] has been applied to other families of sequences [20], the resulting asymptotic merit factor remains 6.

While the merit factor problem originated in communications engineering, both its variants A and B have also attracted interest in theoretical physics [19] and theoretical chemistry [21]. In this context it is considered a notorious problem of combinatorial optimization, usually referred to as the *low autocorrelated binary string problem.* Stochastic search methods such as simulated annealing [22] and evolutionary algorithms [2], [23] have been shown to find sequences with merit factor larger than 8 and occasionally larger than 9 for specific values of $n \leq 200$. However no practical stochastic search method has been demonstrated to produce sequences with merit factor reliably greater than 6 for large n. The difficulty for stochastic search methods seems to be that the combinatorial "landscape" of the search space has an exceptionally large number of local maxima [24], [25].

The merit factor problem has an equivalent formulation in terms of the value of a complex polynomial on the unit circle and as such has also been studied as a problem in analytic number theory [12], [26]. Let $P_A(z) = \sum_{i=0}^{n-1} a_i z^i$ be the polynomial whose coefficients are the elements of the sequence $A = (a_i)$. The L_α norm $||P(z)||_\alpha$ of the polynomial $P(z)$ on the unit circle is defined, for positive α, by

$$||P(z)||_\alpha = \left(\int_0^1 |P(\exp(2\pi\theta\sqrt{-1}))|^\alpha d\theta \right)^{1/\alpha} \tag{3}$$

and it is straightforward to show [12, p. 122] that

$$||P_A(z)||_4^4 = n^2 + 2\sum_{u=1}^{n-1} [C_A(u)]^2. \tag{4}$$

Therefore the merit factor of the sequence A is related to the L_4 norm of the corresponding polynomial of degree $n-1$ by

$$||P_A(z)||_4^4 = n^2 \left(1 + \frac{1}{F(A)} \right), \tag{5}$$

and a large merit factor corresponds to a small L_4 norm. (Note that [12] mostly deals with polynomials of degree n rather than $n-1$ and so the relation (5) appears in a different form in that reference.) Furthermore, since $\int_0^1 |P_A(\exp(2\pi\theta\sqrt{-1}))|^2 d\theta = \sum_{i=0}^{n-1} a_i^2 = n$, we deduce from (2), (3) and (4) that

$$\int_0^1 \left\{ \left| P_A(\exp(2\pi\theta\sqrt{-1})) \right|^2 - n \right\}^2 d\theta = \frac{n^2}{F(A)}. \tag{6}$$

In communications engineering terms, the left-hand side of (6) measures, in terms of power, how much the amplitude spectrum of the signal corresponding to the sequence (a_i) deviates from its mean value n [22]. Therefore a larger merit factor corresponds to a more uniform distribution of the signal energy over the frequency range.

In this paper we will switch between a binary sequence representation and a polynomial representation according to whichever appears to allow a more straightforward treatment.

3 Notation

In order to describe our construction we introduce the following notation. Given a sequence $A = (a_0, a_1, \ldots, a_{n-1})$ of length n and real numbers r, t, where t lies in the interval $[0,1]$, we will write A_r for the sequence $(b_0, b_1, \ldots, b_{n-1})$ obtained by *rotating* the sequence A by a multiple r of its length:

$$b_i := \quad a_{(i+\lfloor rn \rfloor) \bmod n} \quad \text{for } 0 \le i < n,$$

and will write A^t for the sequence $(b_0, b_1, \ldots, b_{\lfloor tn \rfloor - 1})$ obtained by *truncating* A to a fraction t of its length:

$$b_i := \quad a_i \quad \text{for } 0 \le i < \lfloor tn \rfloor.$$

Given sequences $A = (a_0, a_1, \ldots, a_{n-1})$ of length n and $A' = (a'_0, a'_1, \ldots, a'_{n'-1})$ of length n' we will write $A; A'$ for the sequence $(b_0, b_1, \ldots, b_{n+n'-1})$ given by *appending* A' to A:

$$b_i := \begin{cases} a_i & \text{for } 0 \le i < n \\ a'_{i-n} & \text{for } n \le i < n+n'. \end{cases}$$

We use the standard definition of the *periodic autocorrelation* of a sequence $A = (a_0, a_1, \ldots, a_{n-1})$ at shift u:

$$R_A(u) := \sum_{i=0}^{n-1} a_i a_{(i+u) \bmod n} \qquad \text{for } u = 0, 1, \ldots, n-1, \tag{7}$$

so that

$$R_A(u) = C_A(u) + C_A(n-u) \qquad \text{for } u \ne 0.$$

In this paper we are primarily interested in binary sequences, but the definition of periodic/aperiodic autocorrelation and merit factor, as well as the notation given above, can be applied to any real-valued sequence.

Given a sequence $A = (a_0, a_1, \ldots, a_{n-1})$ of length n, write $S(A)$ for the *sum* $\sum_{i=0}^{n-1} a_i$ of the sequence elements. It follows from (1) that

$$2\sum_{u=1}^{n-1} C_A(u) = [S(A)]^2 - \sum_{i=0}^{n-1} a_i^2, \tag{8}$$

which equals $[S(A)]^2 - n$ when A is a binary sequence.

We define the *Legendre sequence* $X = (x_0, x_1, \ldots, x_{n-1})$ of prime length n by:

$$x_i := \begin{cases} 1 & \text{if } i = 0 \\ \left(\frac{i}{n}\right) & \text{if } i > 0, \end{cases}$$

where $\left(\frac{i}{n}\right)$ is the Legendre symbol (which takes the value 1 if i is a quadratic residue modulo n and the value -1 if not).

We represent the fractional part of a real number $x \geq 0$ by $\{x\} := x - \lfloor x \rfloor$.

4 Appending to $\frac{1}{4}$-rotated Legendre Sequences

Let X be a Legendre sequence of prime length n. Turyn studied the merit factor of the rotated Legendre sequence X_r numerically, as reported in [3]. Based on Turyn's work Golay [3] gave a heuristic argument that

$$\frac{1}{\lim\limits_{n\longrightarrow\infty} F(X_r)} = \begin{cases} \frac{1}{6} + 8(r - \frac{1}{4})^2 & \text{for } 0 \leq r \leq \frac{1}{2} \\ \frac{1}{6} + 8(r - \frac{3}{4})^2 & \text{for } \frac{1}{2} \leq r \leq 1, \end{cases} \tag{9}$$

in accordance with Turyn's calculations. Høholdt and Jensen [4] then gave a rigorous proof that (9) is correct. Although Borwein and Choi [27] subsequently determined the exact, rather than the asymptotic, value of $F(X_r)$, the formula (9) will be sufficient for our present purposes.

(9) implies that the asymptotic merit factor of X_r ranges from a minimum of 1.5 to a maximum of 6 and that the maximum value of 6 is attained by the $\frac{1}{4}$-rotated Legendre sequence $X_{\frac{1}{4}}$. Kirilusha and Narayanaswamy [5] investigated, for sequence lengths n up to about 3000, how many arbitrary elements could be appended to the sequence $X_{\frac{1}{4}}$ without decreasing the merit factor significantly below its value of about 6. They observed that when the elements appended to one end of the sequence were identical to the initial elements at the other end of the sequence there was actually a consistent *increase* in the merit factor.

Following their approach, for any real t lying in the interval $[0, 1]$ we consider appending the first $\lfloor tn \rfloor$ elements of $X_{\frac{1}{4}}$ to itself to form the sequence $X_{\frac{1}{4}}; (X_{\frac{1}{4}})^t$ of length $n + \lfloor tn \rfloor$. Table 1 shows the value of $F(X_{\frac{1}{4}})$ and, for various values of t, the value of $F(X_{\frac{1}{4}}; (X_{\frac{1}{4}})^t)$, for the first ten prime values of n greater than 1000. Table 2 shows similar values for $F(X_{\frac{1}{4}})$ and $F(X_{\frac{1}{4}}; (X_{\frac{1}{4}})^t)$, for n in the region of 30000. In all cases for fixed n and increasing t

(reading across a row of the tables) the value of the merit factor increases to a maximum value between 6.06 and 6.20 and then decreases. To our knowledge this, together with the data in [5], is the first numerical evidence that suggests a systematic construction of sequences whose asymptotic merit factor is greater than 6.

We also remark that the maximum value of the merit factor, as t varies, exhibits considerable variation for consecutive values of n in Table 1 but almost none in Table 2. This effect is only partially explained by the more pronounced variation in the base value $F(X_{\frac{1}{4}})$ at smaller values of n. We believe this is an example of "smooth behaviour" in the merit factor arising only for large n, well beyond the range considered in [5]. We believe that the variations in maximum merit factor reported in [5], which were attributed to the distinction between $n \equiv 1 \pmod 4$ and $n \equiv 3 \pmod 4$ and to the distinction between prepending rather than appending the first $\lfloor tn \rfloor$ elements of $X_{\frac{1}{4}}$ to itself, can likewise be explained as non-smooth effects at smaller values of n. Indeed these variations are not apparent for larger n. We also mention again that expressing the number of appended elements as a multiple of n, rather than a power of n as in [5], is key to obtaining behaviour that appears to persist for large n.

We similarly obtained values for $F(X_{\frac{1}{4}})$, and $F(X_{\frac{1}{4}}; (X_{\frac{1}{4}})^t)$ for $t = 0.01, 0.02, \ldots, 0.10$, for 19 different lengths n in the range 2000–77000. For each n we found the same effect as in Tables 1 and 2 for increasing t, with the value of the merit factor increasing to a maximum value between 6.168 and 6.199 before decreasing.

The optimal value $\hat{t}$ maximizing $F(X_{\frac{1}{4}}; (X_{\frac{1}{4}})^t)$ for a given n can be determined as follows: find the (least) integer j in the range $1 \le j \le n$ for which the merit factor of the sequence $X_{\frac{1}{4}}; (X_{\frac{1}{4}})^{j/n}$ of length $n + j$ takes its maximum value, and then set $\hat{t} = j/n$. In determining this value j it is sufficient to perform a full calculation of aperiodic autocorrelations just once, for the initial sequence $X_{\frac{1}{4}}$, and thereafter simply to update the aperiodic autocorrelations as successive sequence elements are appended. We found the value of $\hat{t}$ in this way for 44 values of n in the range 40–260000 and concluded that, for large n,

$$0.031 \le \hat{t} \le 0.036. \tag{10}$$

Figure 1 shows the variation with n of the resulting maximal merit factor $F(X_{\frac{1}{4}}; (X_{\frac{1}{4}})^{\hat{t}})$ — which appears to reach a value greater than 6.2 — and compares this with $F(X_{\frac{1}{4}})$, which approaches the value 6 in accordance with the asymptotic formula (9). We wish now to establish a theoretical framework for these experimental observations.

5 Appending to r-rotated Legendre Sequences

Figure 1 suggests that we can systematically obtain a merit factor of around 6.2 for large n. In fact we can do better than this numerically by considering rotations of Legendre sequences for values of r other than $\frac{1}{4}$. For r, t lying in the interval [0,1] consider the sequence $X_r; (X_r)^t$ of length $n + \lfloor tn \rfloor$ obtained by appending the first $\lfloor tn \rfloor$ elements of X_r to itself. Let $\hat{t}$ maximize $F(X_r; (X_r)^t)$ for a given r.

Figure 2 shows how the value of $F(X_r; (X_r)^{\hat{t}})$ varies with r, for a large fixed length $n = 259499$. We see that there are certain ranges of r in which it is not possible to improve the merit factor of X_r by appending any truncation of X_r to itself. But, outside these ranges, the largest improvement in the merit factor occurs not at $r = \frac{1}{4}$ and $\frac{3}{4}$ (which, from (9), would give the largest asymptotic merit factor for the unappended sequences) but at $r \simeq 0.22$ and $r \simeq 0.72$.

The rotation value $r = 0.22$ is examined in more detail in Figure 3, using the same values for the length n as in Figure 1. It appears that, for large n, we systematically obtain a merit factor $F(X_{0.22}; (X_{0.22})^{\hat{t}})$ of around 6.34, with $\hat{t}$ satisfying

$$0.055 \leq \hat{t} \leq 0.063. \tag{11}$$

Therefore we broaden our objective to: establish theoretically, for large n, the optimal values $\hat{r}$ and $\hat{t}$ maximizing $F(X_r; (X_r)^t)$ and determine $F(X_{\hat{r}}; (X_{\hat{r}})^{\hat{t}})$ explicitly.

6 Analysis

In this section we consider how the asymptotic merit factor of a set of sequences changes under modification of the sequences. We begin by proving that, for any fixed $\alpha < 1$, we can append up to $\lfloor n^\alpha \rfloor$ of the initial elements of a rotated Legendre sequence of length n to itself without changing the asymptotic merit factor. This is contrary to the speculation of [5] and is implied by taking $t = n^{\alpha-1}$ in the following:

Proposition 6.1. *Let X be a Legendre sequence of prime length n and let $r \in [0,1]$ be fixed. Suppose that $t \in [0,1]$ varies with n so that it satisfies $\lim_{n\longrightarrow\infty} t(\log n)^2 = 0$. Then*

$$\lim_{n\longrightarrow\infty} F(X_r; (X_r)^t) = \lim_{n\longrightarrow\infty} F(X_r).$$

Proof. For any $0 \leq r, t \leq 1$, we write $P(r,t,z)$ for the polynomial whose coefficients are the elements of the sequence $(X_r)^t$. Hence $P(r,1,z) + z^n P(r,t,z)$ is the polynomial whose coefficients are the elements of the sequence $X_r; (X_r)^t$. In view of (1) and (4), we have

$$\|P(r,t,z)\|_4^4 = (\lfloor tn \rfloor)^2 + 2 \sum_{u=1}^{\lfloor tn \rfloor - 1} \left(\sum_{i=0}^{\lfloor tn \rfloor - u - 1} x_{(i+\lfloor rn \rfloor) \bmod n} \, x_{(i+\lfloor rn \rfloor + u) \bmod n} \right)^2.$$

It is known (see for example [28]), as a consequence of Weil's theorem [29] on a generalized Riemann hypothesis for curves over a finite field, that if $n \nmid u$ then for any real numbers U and V and integer m,

$$\left| \sum_{U < i < U+V} \left(\frac{i+m}{n} \right) \left(\frac{i+m+u}{n} \right) \right| \ll n^{\frac{1}{2}} \log n.$$

Since $x_i = \left(\frac{i}{n}\right)$ except when $i = 0$, it follows that

$$\|P(r,t,z)\|_4^4 \ll (\lfloor tn \rfloor)^2 + 2 \sum_{u=0}^{\lfloor tn \rfloor - 1} (n^{\frac{1}{2}} \log n)^2 \sim 2\, tn^2 (\log n)^2.$$

Now using the triangle inequality for norms, we get

$$\begin{aligned} & |\|P(r,1,z) + z^n P(r,t,z)\|_4 - \|P(r,1,z)\|_4| \\ \leq \quad & \|z^n P(r,t,z)\|_4 = \|P(r,t,z)\|_4 \ll 2^{\frac{1}{4}} (tn^2 (\log n)^2)^{\frac{1}{4}} \end{aligned}$$

and hence

$$\|P(r,1,z) + z^n P(r,t,z)\|_4 = \|P(r,1,z)\|_4 + O((tn^2(\log n)^2)^{\frac{1}{4}}). \tag{12}$$

If $\lim_{n \longrightarrow \infty} t(\log n)^2 = 0$ then $O((tn^2(\log n)^2)^{\frac{1}{4}})$ is $o(n^{\frac{1}{2}})$. So taking the fourth power in (12) and using (5) we obtain

$$\lim_{n \longrightarrow \infty} F(X_r; (X_r)^t) = \lim_{n \longrightarrow \infty} F(X_r).$$

□

Proposition 6.1 implies that up to $\lfloor n^\alpha \rfloor$ of the initial elements of a rotated Legendre sequence of length n can be appended to the sequence without changing the asymptotic merit factor, for any fixed $\alpha < 1$. In contrast Kirilusha and Narayanaswamy [5] showed that up to $o(\sqrt{n})$ *arbitrary* elements taking the value 1 or -1 can be appended to any binary sequence of length n without changing the asymptotic merit factor. We now prove a more general version of their result, which will be useful in our analysis.

Lemma 6.2. *Let $\{P_n(z)\}$ and $\{Q_n(z)\}$ be sequences of polynomials. Suppose that*

$$\|(P_n - Q_n)(z)\|_\infty = o(\sqrt{n}),$$

where $\|(P_n - Q_n)(z)\|_\infty := \sup_{|z|=1} |(P_n - Q_n)(z)|$. Then

$$\|P_n(z)\|_4 = \|Q_n(z)\|_4 + o(\sqrt{n}).$$

Proof. From the triangle inequality for norms,

$$|\|P_n(z)\|_4 - \|Q_n(z)\|_4| \leq \|(P_n - Q_n)(z)\|_4, \tag{13}$$

while from (3),

$$\|(P_n - Q_n)(z)\|_4 \leq \|(P_n - Q_n)(z)\|_\infty. \tag{14}$$

The result is given by combining (13) and (14). □

Corollary 6.3. *Let $\{A_n\}$ and $\{B_n\}$ be sets of real-valued sequences, where each of A_n and B_n has length n. Suppose that, for each n, the number of non-zero elements of B_n is $o(\sqrt{n})$ and that all the non-zero elements of B_n are bounded in magnitude by a constant independent of n. Then, as $n \longrightarrow \infty$, the elementwise sequence sums $\{A_n + B_n\}$ satisfy*

$$\frac{1}{F(A_n + B_n)} = \frac{1}{F(A_n)} + o(1).$$

Proof. Let $P_n(z)$ and $Q_n(z)$ be the polynomials whose coefficients are the elements of the sequences $A_n + B_n$ and A_n respectively. From Lemma 6.2 we get $\|P_n(z)\|_4 = \|Q_n(z)\|_4 + o(\sqrt{n})$. Take the fourth power and use (5) to obtain the result. □

In the remainder of this section we use Corollary 6.3 to derive an asymptotic expression relating the merit factor of the appended sequence $X_r; (X_r)^t$ to the merit factor of the truncated sequences $(X_r)^t$ and $(X_{r+t})^{1-t}$.

Theorem 6.4. *Let X be a Legendre sequence of prime length n. Let r lie in the interval $[0,1]$ and t lie in the interval $(0,1)$. Then for large n,*

$$\frac{1}{F(X_r;(X_r)^t)} \sim 2\left(\frac{t}{1+t}\right)^2\left(\frac{1}{F((X_r)^t)}+1\right)+\left(\frac{1-t}{1+t}\right)^2\left(\frac{1}{F((X_{r+t})^{1-t})}\right).$$

Proof. We give the proof in detail for the case $n \equiv 3 \pmod 4$, and then indicate how to modify it for the case $n \equiv 1 \pmod 4$.

Set $j = \lfloor tn \rfloor$. For any sequence $A = (a_0, a_1, \ldots, a_{n-1})$ of length n, by definition A^t is the sequence $(a_0, a_1, \ldots, a_{j-1})$ comprising the first j elements of A. Write A' for the sequence $(a_j, a_{j+1}, \ldots, a_{n-1})$ comprising the last $n-j$ elements of A, so that $A = A^t; A'$. It follows from (1) and (7) that we can express the aperiodic autocorrelations of the appended sequence $A; A^t$ in terms of the periodic or aperiodic autocorrelations of the sequences A, A^t and A':

$$C_{A;A^t}(u) = \begin{cases} R_A(u) + C_{A^t}(u) & \text{for } 1 \le u < j \\ R_A(j) & \text{for } u = j \\ R_A(u) - C_{A'}(n-u) & \text{for } j < u < n \\ j & \text{for } u = n \\ C_{A^t}(u-n) & \text{for } n < u < n+j. \end{cases} \tag{15}$$

It is well known (see for example [30, Thm. 6.16]) that, for $n \equiv 3 \pmod 4$, the periodic autocorrelations of $X = (x_0, x_1, \ldots, x_{n-1})$ satisfy

$$R_X(u) = -1 \quad \text{for all } u \neq 0,$$

and therefore

$$R_{X_r}(u) = -1 \quad \text{for all } u \neq 0 \tag{16}$$

since rotation of X does not affect its periodic autocorrelations. So from (2) we have

$$\frac{(n+j)^2}{2F(X_r;(X_r)^t)} = \sum_{u=1}^{n+j-1} [C_{X_r;(X_r)^t}(u)]^2. \tag{17}$$

Now set $A = X_r$ in (15) and use (16) and (17) to deduce that

$$\begin{aligned} \frac{(n+j)^2}{2F(X_r;(X_r)^t)} &= \sum_{u=1}^{j-1}[C_{(X_r)^t}(u)-1]^2 + 1 + \sum_{u=j+1}^{n-1}[C_{(X_r)'}(n-u)+1]^2 + j^2 + \\ &\quad \sum_{u=n+1}^{n+j-1}[C_{(X_r)^t}(u-n)]^2. \end{aligned} \tag{18}$$

We next examine the three summations of (18) in turn. (Note that for large n we can assume that $2 \le j \le n-2$, so that each of the three summation ranges in (18) is non-empty.) From (2) and (8) we have

$$\sum_{u=1}^{j-1}[C_{(X_r)^t}(u)-1]^2 = \frac{j^2}{2F((X_r)^t)} - ([S((X_r)^t)]^2 - j) + (j-1)$$

and

$$\begin{aligned}\sum_{u=j+1}^{n-1}[C_{(X_r)'}(n-u)+1]^2 &= \sum_{u=1}^{n-j-1}[C_{(X_r)'}(u)+1]^2 \\ &= \frac{(n-j)^2}{2F((X_r)')} + ([S((X_r)')]^2 - n + j) + (n-j-1),\end{aligned}$$

and

$$\begin{aligned}\sum_{u=n+1}^{n+j-1}[C_{(X_r)^t}(u-n)]^2 &= \sum_{u=1}^{j-1}[C_{(X_r)^t}(u)]^2 \\ &= \frac{j^2}{2F((X_r)^t)}.\end{aligned}$$

Substitution for these three terms in (18) gives

$$\frac{(n+j)^2}{2F(X_r;(X_r)^t)} = \frac{j^2}{F((X_r)^t)} + \frac{(n-j)^2}{2F((X_r)')} + [S((X_r)')]^2 - [S((X_r)^t)]^2 + j^2 + 2j - 1. \quad (19)$$

Now consider the behaviour of (19) for large n (and fixed r and t). Since $j \sim tn$ we get $j^2+2j-1 \sim t^2n^2$. Since $X_r = (X_r)^t;(X_r)'$ we have $S((X_r)')+S((X_r)^t) = S(X_r) = S(X)$, and it is well known [30] that $S(X) = 1$ for any Legendre sequence X. Therefore

$$[S((X_r)')]^2 - [S((X_r)^t)]^2 = S((X_r)') - S((X_r)^t),$$

which has magnitude at most $(n-j)+j = n$ and so can be neglected in (19). Also, by comparing

$$(X_r)' \equiv (x_{(\lfloor rn\rfloor+j) \bmod n}, \ldots, x_{(\lfloor rn\rfloor-1) \bmod n})$$

with

$$(X_{r+t})^{1-t} \equiv (x_{(\lfloor (r+t)n\rfloor) \bmod n}, \ldots, x_{(\lfloor (r+t)n\rfloor+\lfloor (1-t)n\rfloor-1) \bmod n}),$$

we see that we can write

$$(X_r)' = \begin{cases} (X_{r+t})^{1-t} & \text{if } \{tn\} = 0 \\ (X_{r+t})^{1-t}; x_{(\lfloor rn\rfloor-1) \bmod n} & \text{if } \{tn\} > 0 \text{ and } \{rn\}+\{tn\} < 1 \\ x_{(\lfloor rn\rfloor+j) \bmod n}; (X_{r+t})^{1-t} & \text{if } \{tn\} > 0 \text{ and } 1 \le \{rn\}+\{tn\} < 2. \end{cases}$$

Therefore from Corollary 6.3 we can replace $1/F((X_r)')$ by $1/F((X_{r+t})^{1-t})$ in (19), for large n. Hence from (19) we obtain the asymptotic relationship

$$\frac{(1+t)^2n^2}{2F(X_r;(X_r)^t)} \sim \frac{t^2n^2}{F((X_r)^t)} + \frac{(1-t)^2n^2}{2F((X_{r+t})^{1-t})} + t^2n^2,$$

as required.

We now outline the case $n \equiv 1 \pmod 4$. Consider the sequence $Y = (0, x_1, x_2, \ldots, x_{n-1})$, which differs from the Legendre sequence X in just one element. It is well known [30] that the periodic autocorrelations of Y satisfy

$$R_Y(u) = -1 \quad \text{for all } u \neq 0,$$

and that $S(Y) = 0$. The above proof then carries through (with some modifications to lower order terms) to give

$$\frac{(1+t)^2}{2F(Y_r;(Y_r)^t)} \sim t^2\left(\frac{1}{F((Y_r)^t)} + 1\right) + \frac{(1-t)^2}{2F((Y_{r+t})^{1-t})}.$$

But $(Y_r)^t$ and $(Y_{r+t})^{1-t}$ differ from $(X_r)^t$ and $(X_{r+t})^{1-t}$ respectively in at most one element, and $Y_r;(Y_r)^t$ differs from $X_r;(X_r)^t$ in at most two elements. The result follows from Corollary 6.3. □

The value $t = 1$ is excluded from Theorem 6.4 but a similar argument to that used in its proof shows that, for large n and for r lying in the interval $[0, 1]$,

$$\frac{1}{F(X_r;X_r)} \sim \frac{1}{2}\left(\frac{1}{F(X_r)} + 1\right). \tag{20}$$

7 Asymptotic Merit Factor of Truncated Rotated Legendre Sequences

Our objective is still to determine, for large n, the optimal values $\hat{r}$ and $\hat{t}$ maximizing $F(X_r;(X_r)^t)$ and to determine $F(X_{\hat{r}};(X_{\hat{r}})^{\hat{t}})$ explicitly. We define a region

$$D := \{(r,t) : 0 \leq r \leq 1 \text{ and } 0 < t \leq 1\}.$$

In view of Theorem 6.4 and (20) it is sufficient to determine an asymptotic form (for large n) for the function $1/F((X_r)^t)$ for any $(r,t) \in D$, in other words for the reciprocal merit factor of any non-trivial truncation of any rotation of a Legendre sequence, for large n. We already know the asymptotic form for $t = 1$, from (9).

We firstly investigate how $1/F((X_r)^t)$ behaves for fixed r and t as n grows large, and specifically whether this value converges. Tables 3 and 4 show the value of $1/F((X_r)^t)$ for $r = 0$ and $\frac{1}{2}$ respectively, over a representative range of values of t and a wide range of lengths n. We also made corresponding tables for $r = \frac{1}{4}$ and $\frac{3}{4}$. In all cases it appears that $1/F((X_r)^t)$ approaches a fixed value as n grows large (reading down any column of the tables), although more slowly for smaller values of t. On the basis of this evidence, we shall assume temporarily that the value of $1/F((X_r)^t)$ always converges:

Working Assumption 7.1. *Let X be a Legendre sequence of prime length n. Then the following is well-defined for any $(r,t) \in D$:*

$$f(r,t) := \lim_{n \longrightarrow \infty} \left(\frac{1}{F((X_r)^t)} \right). \tag{21}$$

We next investigate the behaviour of the function $f(r,t)$ by examining numerical data for a representative very large value of n. Accordingly Table 5 shows the variation of $1/F((X_r)^t)$ with r and t for $n = 4433701$. We chose to take r and t each to be multiples of a reciprocal power of 2 because preliminary experiments suggested that the function f exhibits important behaviour at these points. Table 5 is a summary of a more detailed data set in which the values of r and t each vary as multiples of $\frac{1}{64}$.

A property of $f(r,t)$ immediately suggested by the data of Table 5 is the following "reflection" result, which we prove using Corollary 6.3:

Proposition 7.2. *Suppose that $f(r,t)$ is well-defined in (21) for a given $(r,t) \in D$. Then*

$$f(r,t) = f(1-r-t,t).$$

Proof. We suppose that $r + t < 1$. The case $1 \leq r + t < 2$ can be treated similarly. Since

$$(1-r-t)n = (n - \lfloor rn \rfloor - \lfloor tn \rfloor) - (\{rn\} + \{tn\}),$$

we have $\lfloor (1-r-t)n \rfloor = n - k - \lfloor rn \rfloor - \lfloor tn \rfloor$, where

$$k := \begin{cases} 0 & \text{if } \{rn\} + \{tn\} = 0 \\ 1 & \text{if } 0 < \{rn\} + \{tn\} \leq 1 \\ 2 & \text{if } 1 < \{rn\} + \{tn\} < 2. \end{cases}$$

It follows that

$$\left(\frac{\lfloor (1-r-t)n \rfloor + i}{n} \right) = \left(\frac{-1}{n} \right) \left(\frac{\lfloor rn \rfloor + \lfloor tn \rfloor + k - i}{n} \right) \quad \text{for } 0 \leq i \leq n-1.$$

Hence if we write A^* for the sequence A reversed then

$$(X_{1-r-t})^t = \left(\frac{-1}{n} \right) \left((X_{r+\frac{k+1}{n}})^t \right)^*$$

and so

$$\frac{1}{F((X_{1-r-t})^t)} = \frac{1}{F((X_{r+\frac{k+1}{n}})^t)}, \tag{22}$$

since reversal of a sequence does not change its merit factor.

Writing $X = (x_0, x_1, \ldots, x_{n-1})$, we see that

$$(X_r)^t \equiv (x_{\lfloor rn \rfloor}, \ldots, x_{(\lfloor rn \rfloor + k) \bmod n}); Y$$

and

$$(X_{r+\frac{k+1}{n}})^t \equiv Y; (x_{(\lfloor rn \rfloor + \lfloor tn \rfloor) \bmod n}, \ldots, x_{(\lfloor rn \rfloor + \lfloor tn \rfloor + k) \bmod n}),$$

where the sequences $(X_r)^t$ and $(X_{r+\frac{k+1}{n}})^t$ share the common subsequence

$$Y := (x_{\lfloor rn \rfloor + k + 1}, \ldots, x_{\lfloor rn \rfloor + \lfloor tn \rfloor - 1}).$$

Since $k \leq 2$, the result follows using (22) and Corollary 6.3. □

Borwein and Choi [27] determined that

$$f(0, \tfrac{1}{2}) \text{ is well-defined and equals } \tfrac{1}{3}, \tag{23}$$

but as far as we are aware no values of $f(r,t)$ have previously been determined for other $(r,t) \in D$ apart from the value of $f(r,1)$, which is given by (9). We can use (23) and Proposition 7.2 to obtain the following result that does not depend on any assumption of convergence:

Corollary 7.3. *Let X be a Legendre sequence of prime length n. Then*

$$\lim_{n \longrightarrow \infty} F(X; X^{\frac{1}{2}}) = 3.$$

Proof. From (23) and Proposition 7.2 we have $f(\frac{1}{2}, \frac{1}{2}) = f(0, \frac{1}{2}) = \frac{1}{3}$. Apply Theorem 6.4 with $r = 0$ and $t = \frac{1}{2}$. □

Having established Proposition 7.2 we carefully studied the data set underlying Table 5 in order to postulate an explicit formula for the value of $f(r,t)$ for any $(r,t) \in D$. We noted initially that $1/F((X_r)^t) \simeq 1/F((X_{r+\frac{1}{2}})^t)$ for all (r,t), from which we concluded that $f(r,t) = f(r + \frac{1}{2}, t)$. We then observed that certain rational values for $f(r,t)$ were strongly suggested by the data, for example $f(0, \frac{3}{4}) = f(\frac{1}{4}, \frac{3}{4}) = 4/9$ and $f(\frac{3}{8}, \frac{3}{4}) = 2/9$. Having fixed several of these key values we found that, for each constant t, the complete data for $0 \leq r \leq \frac{1}{2}$ could be fitted very well by three quadratic functions in r. By comparing several such fitted functions for different values of t, in particular the values $t = \frac{1}{4}$, $\frac{1}{2}$ and $\frac{3}{4}$, we reached the following working assumption:

Working Assumption 7.4. *Working Assumption 7.1 holds and $f(r,t)$ is given for any $(r,t) \in D$ by:*

$$f(r + \tfrac{1}{2}, t) = f(r,t) \quad \text{for } 0 \leq r \leq \tfrac{1}{2} \text{ and } 0 < t \leq 1 \tag{24}$$

and, for $0 < t \leq \frac{1}{2}$,

$$f(r,t) \;=\; \begin{cases} 1 - \frac{4}{3}t & \text{for } 0 \leq r \leq \frac{1}{2} - t \\ 1 - \frac{4}{3}t + \dfrac{4(r - \frac{1}{2} + t)^2}{t^2} & \text{for } \frac{1}{2} - t \leq r \leq \frac{1-t}{2} \\ 1 - \frac{4}{3}t + \dfrac{4(r - \frac{1}{2})^2}{t^2} & \text{for } \frac{1-t}{2} \leq r \leq \frac{1}{2} \end{cases} \tag{25}$$

and, for $\frac{1}{2} < t \leq 1$,

$$f(r,t) \quad = \quad \begin{cases} 1 - \frac{4}{3}t + \dfrac{4(r - \frac{1}{2} + t)^2}{t^2} & \text{for } 0 \leq r \leq \frac{1-t}{2} \\ 1 - \frac{4}{3}t + \dfrac{4(r - \frac{1}{2})^2}{t^2} & \text{for } \frac{1-t}{2} \leq r \leq 1 - t \\ 1 - \frac{4}{3}t + \dfrac{8(r - \frac{3}{4} + \frac{t}{2})^2 + 2(t - \frac{1}{2})^2}{t^2} & \text{for } 1 - t \leq r \leq \frac{1}{2}. \end{cases}$$

The proposed function $f(r,t)$ specified in Working Assumption 7.4 is illustrated in Figure 4. We see from the graphs and from (24) and (25) that, for $t \leq \frac{1}{2}$, the function contains two "spikes" of width t and height 1, at $r = \frac{1-t}{2}$ and $r = 1 - \frac{t}{2}$. We wish however to avoid any possible problems of discontinuity in $f(r,t)$ as $t \longrightarrow 0$. To do so, observe that the function $t^2 f(r,t)$ given by (25) does not contain any such spikes. Furthermore note from Theorem 6.4 that it is sufficient for our purposes to know the asymptotic behaviour of $t^2/F((X_r)^t)$ rather than that of $1/F((X_r)^t)$. We therefore discard Working Assumptions 7.1 and 7.4 and replace them with the following conjecture:

Conjecture 7.5. *Let X be a Legendre sequence of prime length n. Then*

$$g(r,t) := \begin{cases} \lim\limits_{n \longrightarrow \infty} \left(\dfrac{t^2}{F((X_r)^t)} \right) & \text{for } 0 < t \leq 1 \\ 0 & \text{for } t = 0 \end{cases} \tag{26}$$

is well-defined for any $r, t \in [0,1]$ and is given by

$$g(r,t) = t^2(1 - \tfrac{4}{3}t) + h(r,t),$$

where:

$$h(r + \tfrac{1}{2}, t) := h(r,t) \quad \text{for } 0 \leq r \leq \tfrac{1}{2} \text{ and } 0 \leq t \leq 1$$

and, for $0 \leq t \leq \frac{1}{2}$,

$$h(r,t) \quad := \quad \begin{cases} 0 & \text{for } 0 \leq r \leq \frac{1}{2} - t \\ 4(r - \frac{1}{2} + t)^2 & \text{for } \frac{1}{2} - t \leq r \leq \frac{1-t}{2} \\ 4(r - \frac{1}{2})^2 & \text{for } \frac{1-t}{2} \leq r \leq \frac{1}{2} \end{cases} \tag{27}$$

and, for $\frac{1}{2} < t \leq 1$,

$$h(r,t) \quad := \quad \begin{cases} 4(r - \frac{1}{2} + t)^2 & \text{for } 0 \leq r \leq \frac{1-t}{2} \\ 4(r - \frac{1}{2})^2 & \text{for } \frac{1-t}{2} \leq r \leq 1 - t \\ 8(r - \frac{3}{4} + \frac{t}{2})^2 + 2(t - \frac{1}{2})^2 & \text{for } 1 - t \leq r \leq \frac{1}{2}. \end{cases} \tag{28}$$

The proposed function $g(r,t)$ specified in Conjecture 7.5 is illustrated in Figure 5. Figure 6 shows how well this proposed function $g(r,t)$ fits the data values at increasing lengths n by plotting the discrepancy

$$d(r,t) := \frac{t^2}{F((X_r)^t)} - g(r,t)$$

for

$$(r,t) \in G := \{0, 1/64, 2/64, \ldots, 1\} \times \{1/64, 2/64, \ldots, 1\}.$$

It is visually striking that the fit improves with increasing n, and indeed the data underlying Figure 6 show that

$$\max_{(r,t)\in G} \mid d(r,t) \mid = \begin{cases} 0.00484 & \text{for } n = 22783 \\ 0.00122 & \text{for } n = 259499 \\ 0.00025 & \text{for } n = 4433701. \end{cases} \tag{29}$$

Although we were unable to find a theoretical proof of Conjecture 7.5, we regard Figure 6 and (29) as compelling evidence in its favour.

We can prove a similar reflection result to that of Proposition 7.2 (using the same proof but simply multiplying (22) by t^2 before applying Corollary 6.3):

Proposition 7.6. *Suppose that $g(r,t)$ is well-defined in* (26) *for given $r,t \in [0,1]$. Then*

$$g(r,t) = g(1-r-t,t)$$

and

$$h(r,t) = h(1-r-t,t).$$

Proposition 7.6 shows that, for any fixed t and for r restricted to the range $[0, \frac{1}{2}]$, the function $h(r,t)$ is symmetrical about $r = \frac{1-t}{2}$. Therefore the following definitions for $h(r,t)$ are superfluous in Conjecture 7.5: for the range $\frac{1-t}{2} \le r \le \frac{1}{2}$ in (27) and for the range $0 \le r \le \frac{1-t}{2}$ in (28).

8 Consequences of Conjecture 7.5

In the previous section we presented evidence in favour of the asymptotic form $g(r,t)$ for $t^2/F((X_r)^t)$ specified in Conjecture 7.5. Assuming this form to be correct, we can combine the results of Theorem 6.4 with (9) and (20) to express the asymptotic merit factor of the appended sequence $X_r\,;(X_r)^t$ explicitly:

Corollary 8.1. *Subject to Conjecture* 7.5,

$$\lim_{n\longrightarrow\infty} F(X_r\,;(X_r)^t) = \frac{(1+t)^2}{2(g(r,t)+t^2)+g(r+t,1-t)} \quad \textit{for } r,t \in [0,1]. \tag{30}$$

Let $\hat{t}$ be the value of $t \in [0,1]$ that maximizes the right hand side of (30) for a given $r \in [0,1]$. Figures 7 shows how $\hat{t}$, determined numerically in this way, varies with r. Figure 8 shows the corresponding variation of $\lim_{n\longrightarrow\infty} F(X_r;(X_r)^{\hat{t}})$ with r, subject to Conjecture 7.5, and compares it with the value of $\max_{t\in[0,1]} F(X_r;(X_r)^t)$ for $n = 259499$ previously displayed in Figure 2. There is excellent agreement, for all r, between the predicted asymptotic value and the data value at length $n = 259499$. This provides further evidence in favour of Conjecture 7.5.

We now examine two cases of particular interest in detail: the globally optimal value $\hat{r}$, and the value $r = \frac{1}{4}$ that was the starting point for our investigation.

Corollary 8.2. *Let X be a Legendre sequence of prime length n and assume Conjecture 7.5 to be correct. Then the maximum of $\lim_{n\longrightarrow\infty} F(X_r;(X_r)^t)$ over $r,t \in [0,1]$ is given by*

$$\begin{aligned} \hat{r} &\simeq 0.2211 \text{ or } 0.7211, \\ \hat{t} &\simeq 0.0578, \\ \lim_{n\longrightarrow\infty} F(X_{\hat{r}};(X_{\hat{r}})^{\hat{t}}) &\simeq 6.3421. \end{aligned}$$

Furthermore the maximum of $\lim_{n\longrightarrow\infty} F(X_{\frac{1}{4}};(X_{\frac{1}{4}})^t)$ over $t \in [0,1]$ is given by

$$\begin{aligned} \hat{t} &\simeq 0.0338, \\ \lim_{n\longrightarrow\infty} F(X_{\frac{1}{4}};(X_{\frac{1}{4}})^{\hat{t}}) &\simeq 6.2018. \end{aligned}$$

Proof. We shall find optimal values $\hat{r}$ and $\hat{t}$ in the region $0 \le r+t \le \frac{1}{2}$; other regions for $r,t \in [0,1]$ can be handled similarly. In this region, by Conjecture 7.5 we have

$$\begin{aligned} g(r,t) &= t^2(1-\tfrac{4}{3}t), \\ g(r+t,1-t) &= (1-t)^2(\tfrac{4}{3}t-\tfrac{1}{3}) + 8(r-\tfrac{1}{4}+\tfrac{t}{2})^2 + 2(t-\tfrac{1}{2})^2. \end{aligned}$$

Substitution in (30) gives

$$\lim_{n\longrightarrow\infty} F(X_r;(X_r)^t) = \frac{6(1+t)^2}{-8t^3+18t^2+1+48(r-\frac{1}{4}+\frac{t}{2})^2}. \tag{31}$$

This function is clearly maximized at

$$\hat{r} = \tfrac{1}{4} - \tfrac{t}{2}, \tag{32}$$

and, by differentiation, the maximum value of (31) occurs when t satisfies

$$4t^3 + 12t^2 - 18t + 1 = 0.$$

Solving this cubic equation numerically for t we find $\hat{t} \simeq 0.0578$, and substitution in (31) then gives $\lim_{n\longrightarrow\infty} F(X_{\hat{r}};(X_{\hat{r}})^{\hat{t}}) \simeq 6.3421$. From (32) we get $\hat{r} \simeq 0.2211$.

If instead we fix $r = \frac{1}{4}$ then the maximum value of (31) occurs when t satisfies

$$4t^3 + 12t^2 - 30t + 1 = 0,$$

which has the solution $\hat{t} \simeq 0.0338$. From (31) we deduce that $\lim_{n\longrightarrow\infty} F(X_{\frac{1}{4}}; (X_{\frac{1}{4}})^{\hat{t}}) \simeq 6.2018$. □

The numerical values of Corollary 8.2 accord well with the data of (11) and Figure 3 for $r = 0.22$, and of (10) and Figure 1 for $r = \frac{1}{4}$.

9 Modified Jacobi Sequences

A *modified Jacobi sequence* $Y = (y_0, y_1, \ldots, y_{n-1})$ of length $n = pq$, for p and q distinct primes, is defined in [20] by:

$$y_i := \begin{cases} 1 & \text{if } i \equiv 0 \pmod{q} \\ -1 & \text{if } i > 0 \text{ and } i \equiv 0 \pmod{p} \\ \left(\frac{i}{n}\right) & \text{otherwise.} \end{cases}$$

The case $q = p + 2$ is equivalent to a Twin Prime sequence. Jensen, Høholdt and Jensen showed in [20] that, provided p and q both grow roughly as fast as each other,

$$\lim_{n\longrightarrow\infty} F(Y_r) = \lim_{n\longrightarrow\infty} F(X_r)$$

as given in (9). In particular the asymptotic merit factor of a $\frac{1}{4}$-rotated modified Jacobi sequence is 6. It was noted in [20] that convergence of the merit factor to its asymptotic value appears to be faster for $q \equiv p \pmod 4$ than for $q \equiv p + 2 \pmod 4$.

It is natural to ask how the merit factor of a modified Jacobi sequence changes when the initial elements are appended to itself. We formed the graph corresponding to Figure 2 for a modified Jacobi sequence of length $n = 272953 = 499 \times 547$ and $n = 272483 = 521 \times 523$. Both graphs are very similar to Figure 2, although in the second graph the peak for the appended sequence occurs at 6.25 rather than 6.34. We also formed the plot corresponding to the case $n = 4433701$ of Figure 6 for a modified Jacobi sequence of length $n = 4447397 = 2087 \times 2131$ and $n = 4460543 = 2111 \times 2113$, using the same function $g(r, t)$ but a coarser grid

$$(r, t) \in G' := \{0, 1/16, 2/16, \ldots, 1\} \times \{1/16, 2/16, \ldots, 1\}.$$

The maximum discrepancy over G' for these values of n has magnitude 0.00016 and 0.00133 respectively. These experiments suggest that, provided p and q grow roughly as fast as each other, appending the initial elements of a modified Jacobi sequence to itself produces the same asymptotic behaviour as for Legendre sequences, with faster convergence when $q \equiv p \pmod 4$.

10 Conclusions

We have found a systematic synthesis for binary sequences yielding a merit factor greater than 6, for sequence lengths well in excess of one million. We believe this meets the spirit of Golay's 1983 challenge [3] as quoted in Section 1.

We believe that the evidence presented here provides good reason to reject the 1988 conjecture of Høholdt and Jensen [4] as quoted in Section 1. Although our theoretical results depend on the unproven Conjecture 7.5, we believe that this conjecture is strongly supported by the evidence of Figures 6 and 8. In fact we regard the true maximal asymptotic merit factor of binary sequences to be a completely open problem now. We certainly do not see any reason to believe that the value of approximately 6.3421, derived in Corollary 8.2, is the true value of $\limsup_{n\longrightarrow\infty} F_n$.

We conclude with the following open questions:

1. Can we prove Conjecture 7.5 and so establish rigorously that $\limsup_{n\longrightarrow\infty} F_n > 6.34$?

2. Can we construct a family of binary sequences whose asymptotic merit factor appears to be greater than that of the appended rotated Legendre sequences described here?

3. Can we construct a family of binary sequences whose asymptotic merit factor is an integer value greater than 6?

4. Can we prove that $\limsup_{n\longrightarrow\infty} F_n$ is bounded above?

11 Acknowledgements

We would like to thank Jen Katerenchuk for assisting with some Maple programming, and Bill Munro for his generous help in configuring a key part of the infrastructure on which our computational experiments were carried out.

12 Notes Added in Proof

After submission of the original manuscript in May 2003 we were made aware of independent work by Kristiansen [31], also inspired by Kirilusha and Narayanaswamy [5], which presented sequences of length up to 20,000 having merit factor greater than 6.3. Each of the sequences in [31] was obtained by searching over a set of sequences derived from a Legendre sequence. [31] gives an approximate value for the total number of sequence elements resulting from the search but does not contain a theoretical explanation of the merit factor properties of the sequences. In response to a preprint of the current paper, Kristiansen and Parker [32] recognised that the sequences in [31] could more easily be viewed as an appending of a rotated Legendre sequence.

R. Turyn confirmed, in personal communication during 2003, our reading of the historical record as set out in the first paragraph of Section 2.

References

[1] S. Mertens and H. Bauke, "Ground states of the Bernasconi model with open boundary conditions," Web page, <http://odysseus.nat.uni-magdeburg.de/~mertens/bernasconi/open.dat>, accessed in July 2004.

[2] C. de Groot, D. Würtz, and K. Hoffmann, "Low autocorrelation binary sequences: exact enumeration and optimization by evolutionary strategies," *Optimization*, vol. **23**, p. 1992, 369–384.

[3] M. Golay, "The merit factor of Legendre sequences," *IEEE Trans. Inform. Theory*, vol. **IT-29**, pp. 934–936, 1983.

[4] T. Høholdt and H. Jensen, "Determination of the merit factor of Legendre sequences," *IEEE Trans. Inform. Theory*, vol. **34**, pp. 161–164, 1988.

[5] A. Kirilusha and G. Narayanaswamy, "Construction of new asymptotic classes of binary sequences based on existing asymptotic classes," Dept. Math. and Comput. Science, University of Richmond," Summer Science Program Technical Report, July 1999.

[6] R. Turyn, "Sequences with small correlation," in *Error Correcting Codes*, H. Mann, Ed. New York: Wiley, 1968, pp. 195–228.

[7] R. Turyn and J. Storer, "On binary sequences," *Proc. Amer. Math. Soc.*, vol. **12**, pp. 394–399, 1961.

[8] R. Turyn, "Character sums and difference sets," *Pacific J. Math.*, vol. **15**, pp. 319–346, 1965.

[9] T. Beth, D. Jungnickel, and H. Lenz, *Design Theory*, 2nd ed. Cambridge: Cambridge University Press, 1999, volumes I and II.

[10] M. Golay, "A class of finite binary sequences with alternate autocorrelation values equal to zero," *IEEE Trans. Inform. Theory*, vol. **IT-18**, pp. 449–450, 1972.

[11] P. Erdös, "An inequality for the maximum of trigonometric polynomials," *Ann. Polon. Math.*, vol. **12**, pp. 151–154, 1962.

[12] P. Borwein, *Computational Excursions in Analysis and Number Theory*, ser. CMS Books in Mathematics. New York: Springer-Verlag, 2002.

[13] B. Schmidt, "Cyclotomic integers and finite geometry," *J. Am. Math. Soc.*, vol. **12**, pp. 929–952, 1999.

[14] S. Mertens, "Exhaustive search for low-autocorrelation binary sequences," *J. Phys. A: Math. Gen.*, vol. **29**, pp. L473–L481, 1996.

[15] M. Golay, "Sieves for low autocorrelation binary sequences," *IEEE Trans. Inform. Theory*, vol. **IT-23**, pp. 43–51, 1977.

[16] M. Golay and D. Harris, "A new search for skewsymmetric binary sequences with optimal merit factors," *IEEE Trans. Inform. Theory*, vol. **36**, pp. 1163–1166, 1990.

[17] J. Knauer, "Merit factor records," Web page, `<http://www.cecm.sfu.ca/~jknauer/labs/records.html>`, accessed in July 2004.

[18] M. Golay, "The merit factor of long low autocorrelation binary sequences," *IEEE Trans. Inform. Theory*, vol. **IT-28**, pp. 543–549, 1982.

[19] J. Bernasconi, "Low autocorrelation binary sequences: statistical mechanics and configuration state analysis," *J. Physique*, vol. **48**, pp. 559–567, 1987.

[20] J. Jensen, H. Jensen, and T. Høholdt, "The merit factor of binary sequences related to difference sets," *IEEE Trans. Inform. Theory*, vol. **37**, pp. 617–626, 1991.

[21] P. Stadler, "Landscapes and their correlation functions," *J. Math. Chem.*, vol. **20**, pp. 1–45, 1996.

[22] G. Beenker, T. Claasen, and P. Hermens, "Binary sequences with a maximally flat amplitude spectrum," *Philips J. Res.*, vol. **40**, pp. 289–304, 1985.

[23] B. Militzer, M. Zamparelli, and D. Beule, "Evolutionary search for low autocorrelated binary sequences," *IEEE Trans. Evol. Comput.*, vol. **2**, pp. 34–39, 1998.

[24] V. de Oliveira, J. Fontanari, and P. Stadler, "Metastable states in high order short-range spin glasses," *J. Phys. A: Math. Gen.*, vol. **32**, pp. 8793–8802, 1999.

[25] F. Ferreira, J. Fontanari, and P. Stadler, "Landscape statstics of the low autocorrelated binary string problem," *J. Phys. A: Math. Gen.*, vol. **33**, pp. 8635–8647, 2000.

[26] D. Newman and J. Byrnes, "The L^4 norm of a polynomial with coefficients ± 1," *Amer. Math. Monthly*, vol. **97**, pp. 42–45, 1990.

[27] P. Borwein and K.-K. Choi, "Explicit merit factor formulae for Fekete and Turyn polynomials," *Trans. Amer. Math. Soc.*, vol. **354**, pp. 219–234, 2002.

[28] C. Mauduit and A. Sárközy, "On finite pseudorandom binary sequences I: Measure of pseudorandomness, the Legendre symbol," *Acta Arith.*, vol. **82**, pp. 365–377, 1997.

[29] A. Weil, *Sur les courbes algébriques et les variétés qui s'en déduisent*, ser. Act. Sci. Ind. 1041. Paris: Hermann, 1948.

[30] T. Apostol, *Introduction to Analytic Number Theory*, ser. UTM. New York: Springer-Verlag, 1976.

[31] R. Kristiansen, "On the aperiodic autocorrelation of binary sequences," Master's thesis, University of Bergen, March 2003.

[32] R. Kristiansen and M. Parker, "Binary sequences with merit factor > 6.3," *IEEE Trans. Inform. Theory*, to appear.

		t										
		0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10
	1009	5.88	5.98	6.02	6.06	6.02	5.98	5.90	5.86	5.77	5.65	5.52
	1013	5.86	5.98	6.07	6.11	6.11	6.05	6.02	5.94	5.77	5.66	5.53
	1019	5.99	6.08	6.16	6.17	6.14	6.09	6.03	5.94	5.88	5.78	5.63
	1021	5.85	5.95	6.02	6.08	6.06	6.01	6.00	5.90	5.76	5.63	5.53
	1031	5.93	6.04	6.09	6.13	6.13	6.11	6.06	5.93	5.84	5.72	5.57
n	1033	5.89	6.00	6.06	6.08	6.09	6.07	6.03	5.89	5.78	5.67	5.50
	1039	5.91	6.02	6.09	6.13	6.13	6.09	6.06	6.01	5.88	5.73	5.61
	1049	5.92	5.99	6.04	6.08	6.08	6.03	5.98	5.90	5.81	5.66	5.55
	1051	5.92	6.02	6.08	6.12	6.14	6.11	6.02	5.94	5.83	5.74	5.61
	1061	5.97	6.09	6.18	6.20	6.21	6.13	6.07	5.99	5.88	5.74	5.59

Table 1: Merit factor of $\frac{1}{4}$-rotated Legendre sequence with length n of around 1000, after appending the first $\lfloor tn \rfloor$ sequence elements to itself

		t										
		0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10
	30011	6.00	6.10	6.17	6.20	6.19	6.16	6.10	6.01	5.89	5.76	5.62
	30013	6.00	6.10	6.17	6.20	6.19	6.16	6.09	6.01	5.90	5.76	5.62
	30029	6.00	6.10	6.17	6.20	6.19	6.16	6.09	6.00	5.89	5.77	5.63
	30047	6.00	6.10	6.17	6.20	6.20	6.16	6.09	6.00	5.90	5.76	5.62
	30059	6.00	6.10	6.17	6.20	6.19	6.15	6.08	5.99	5.88	5.76	5.61
n	30071	6.00	6.10	6.17	6.20	6.20	6.15	6.09	5.99	5.88	5.76	5.62
	30089	6.00	6.10	6.17	6.20	6.20	6.16	6.09	6.00	5.89	5.77	5.63
	30091	6.00	6.10	6.16	6.20	6.20	6.16	6.09	6.01	5.89	5.76	5.63
	30097	5.99	6.10	6.16	6.19	6.18	6.15	6.08	5.99	5.88	5.75	5.61
	30103	6.00	6.10	6.17	6.20	6.20	6.16	6.10	6.01	5.90	5.77	5.63

Table 2: Merit factor of $\frac{1}{4}$-rotated Legendre sequence with length n of around 30000, after appending the first $\lfloor tn \rfloor$ sequence elements to itself

		t									
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
	2003	0.82120	0.77660	0.54698	0.42961	0.33463	0.32866	0.40346	0.49368	0.59069	0.66690
	3001	0.99240	0.77913	0.60592	0.44554	0.33328	0.32149	0.39359	0.49282	0.58852	0.66700
	4507	0.85856	0.78075	0.58090	0.48456	0.33436	0.30393	0.39638	0.49254	0.59033	0.66678
	6761	0.80550	0.71960	0.59985	0.46753	0.33414	0.31108	0.39344	0.49659	0.59091	0.66681
	10133	0.83457	0.72741	0.60279	0.46025	0.33388	0.31414	0.39256	0.49616	0.59047	0.66677
	15193	0.90832	0.74063	0.60008	0.47326	0.33351	0.30867	0.39379	0.49562	0.58975	0.66673
n	22783	0.85412	0.72447	0.59754	0.46222	0.33354	0.31323	0.39371	0.49635	0.59020	0.66660
	34171	0.88902	0.73565	0.60971	0.47253	0.33350	0.30859	0.39152	0.49563	0.58986	0.66669
	51263	0.86484	0.74406	0.59849	0.46755	0.33329	0.31065	0.39345	0.49514	0.59009	0.66656
	76907	0.85874	0.73222	0.59906	0.46671	0.33334	0.31114	0.39330	0.49590	0.59020	0.66665
	115331	0.86442	0.72956	0.60206	0.46565	0.33334	0.31159	0.39282	0.49607	0.59015	0.66666
	172999	0.86444	0.73363	0.60431	0.46595	0.33338	0.31146	0.39241	0.49581	0.59013	0.66666
	259499	0.87862	0.73916	0.60213	0.46666	0.33336	0.31112	0.39281	0.49546	0.58998	0.66667

Table 3: Variation with length n of the reciprocal merit factor $1/F((X_0)^t)$ of the unrotated Legendre sequence over a representative range of truncation fractions t.

		t									
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
	2003	0.71620	0.72435	0.57860	0.45893	0.33638	0.31581	0.39567	0.49601	0.58923	0.66432
	3001	0.93196	0.77596	0.61680	0.46927	0.33430	0.30958	0.38937	0.49275	0.58739	0.66475
	4507	0.83983	0.79637	0.55720	0.46857	0.33359	0.31108	0.40125	0.49088	0.58910	0.66566
	6761	0.89459	0.73324	0.59457	0.45467	0.33320	0.31735	0.39382	0.49569	0.58946	0.66655
	10133	0.84647	0.72068	0.59527	0.46812	0.33314	0.31069	0.39409	0.49652	0.59041	0.66679
	15193	0.87606	0.74556	0.61416	0.47316	0.33298	0.30839	0.39065	0.49527	0.58980	0.66609
n	22783	0.84305	0.75499	0.59303	0.45949	0.33321	0.31426	0.39441	0.49434	0.59028	0.66653
	34171	0.83971	0.72962	0.59754	0.46394	0.33327	0.31234	0.39353	0.49602	0.59028	0.66655
	51263	0.88711	0.73443	0.59552	0.46737	0.33334	0.31094	0.39392	0.49566	0.58981	0.66661
	76907	0.85188	0.73329	0.60012	0.46444	0.33330	0.31203	0.39312	0.49570	0.59015	0.66650
	115331	0.84639	0.72879	0.59946	0.46496	0.33331	0.31178	0.39323	0.49606	0.59030	0.66657
	172999	0.88137	0.73317	0.59999	0.46706	0.33332	0.31091	0.39320	0.49583	0.58992	0.66665
	259499	0.86580	0.73083	0.60299	0.46601	0.33331	0.31141	0.39263	0.49597	0.59011	0.66664

Table 4: Variation with length n of the reciprocal merit factor $1/F((X_{\frac{1}{2}})^t)$ of the $\frac{1}{2}$-rotated Legendre sequence over a representative range of truncation fractions t.

		t							
		1/16	2/16	3/16	4/16	5/16	6/16	7/16	8/16
	0	0.91447	0.83061	0.74879	0.66657	0.58336	0.49954	0.41691	0.33334
	1/16	0.91850	0.83254	0.75034	0.66641	0.58336	0.50032	0.41650	0.39567
	2/16	0.91200	0.83446	0.74946	0.66685	0.58407	0.50029	0.49838	0.58345
	3/16	0.91397	0.83224	0.75078	0.66676	0.58300	0.61125	0.74334	0.89591
	4/16	0.91501	0.83311	0.74956	0.66597	0.74330	0.94501	1.15135	1.33337
	5/16	0.92007	0.83236	0.74908	0.91634	1.22387	1.50036	1.15135	0.89591
	6/16	0.91403	0.83224	1.19486	1.66772	1.22389	0.94502	0.74334	0.58345
	7/16	0.91403	1.82853	1.19488	0.91635	0.74330	0.61126	0.49839	0.39567
r	8/16	0.91403	0.83224	0.74908	0.66597	0.58301	0.50029	0.41650	0.33333
	9/16	0.91402	0.83237	0.74955	0.66675	0.58407	0.50033	0.41691	0.39600
	10/16	0.92007	0.83310	0.75078	0.66685	0.58336	0.49954	0.49855	0.58322
	11/16	0.91502	0.83225	0.74945	0.66640	0.58335	0.61108	0.74290	0.89576
	12/16	0.91398	0.83447	0.75033	0.66657	0.74324	0.94404	1.15105	1.33328
	13/16	0.91198	0.83252	0.74878	0.91713	1.22279	1.49958	1.15105	0.89576
	14/16	0.91849	0.83061	1.19497	1.66408	1.22279	0.94404	0.74290	0.58322
	15/16	0.91448	1.83675	1.19498	0.91713	0.74324	0.61108	0.49855	0.39599
	1	0.91447	0.83061	0.74879	0.66657	0.58336	0.49954	0.41691	0.33334

		t							
		9/16	10/16	11/16	12/16	13/16	14/16	15/16	1
	0	0.29924	0.32683	0.38085	0.44445	0.50845	0.56808	0.62112	0.66667
	1/16	0.44738	0.52668	0.61228	0.69439	0.76871	0.83326	0.62112	0.44792
	2/16	0.69463	0.80681	0.90989	1.00029	0.76871	0.56808	0.40777	0.29167
	3/16	1.04031	1.16682	0.90989	0.69440	0.50845	0.36396	0.26558	0.19792
	4/16	1.04031	0.80681	0.61228	0.44446	0.31902	0.24147	0.19446	0.16667
	5/16	0.69463	0.52668	0.38085	0.27781	0.22439	0.20070	0.19445	0.19792
	6/16	0.44738	0.32683	0.24862	0.22220	0.22432	0.24150	0.26554	0.29167
	7/16	0.29924	0.24655	0.24847	0.27777	0.31906	0.36397	0.40779	0.44792
r	8/16	0.29948	0.32656	0.38092	0.44452	0.50843	0.56805	0.62112	0.66667
	9/16	0.44748	0.52662	0.61226	0.69448	0.76871	0.83343	0.62113	0.44792
	10/16	0.69436	0.80646	0.90967	0.99988	0.76872	0.56805	0.40779	0.29167
	11/16	1.04013	1.16654	0.90967	0.69448	0.50843	0.36397	0.26554	0.19792
	12/16	1.04013	0.80646	0.61227	0.44452	0.31906	0.24150	0.19445	0.16667
	13/16	0.69436	0.52662	0.38092	0.27777	0.22432	0.20070	0.19446	0.19792
	14/16	0.44748	0.32656	0.24847	0.22220	0.22439	0.24147	0.26558	0.29167
	15/16	0.29948	0.24655	0.24862	0.27781	0.31902	0.36396	0.40777	0.44792
	1	0.29924	0.32683	0.38085	0.44445	0.50845	0.56808	0.62112	0.66667

Table 5: Variation of the reciprocal merit factor $1/F$ of $(X_r)^t$ with r and t, where X is a Legendre sequence of length 4433701.

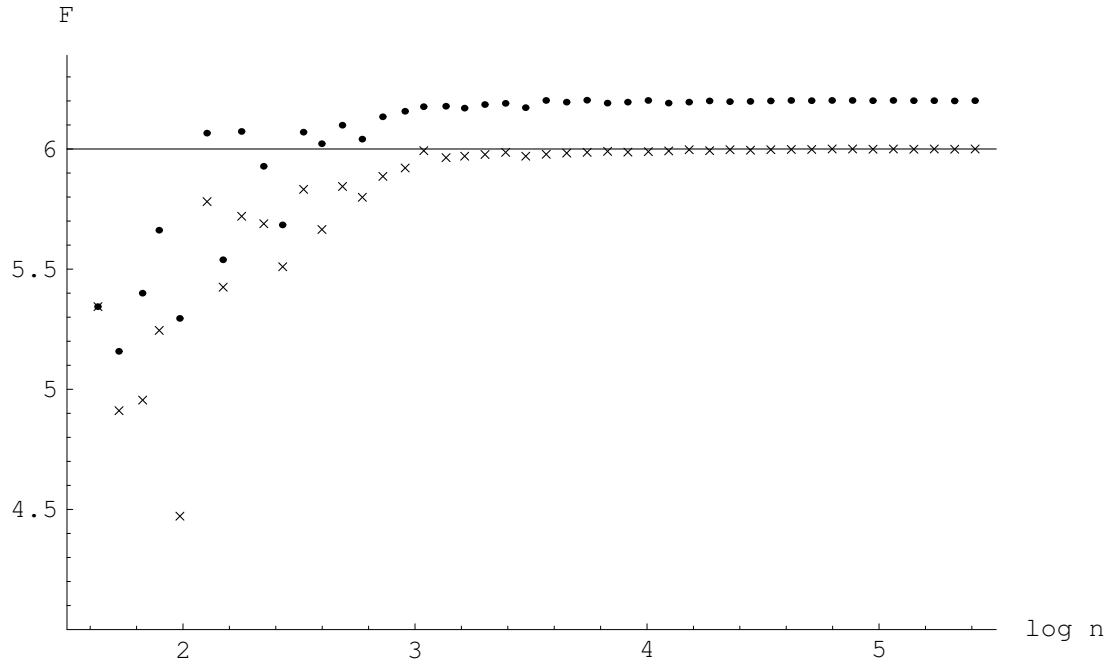


Figure 1: The merit factor of the $\frac{1}{4}$-rotated Legendre sequence of length n before ($\times$) and after ($\bullet$) appending of the optimal number of its own initial elements, for varying n.

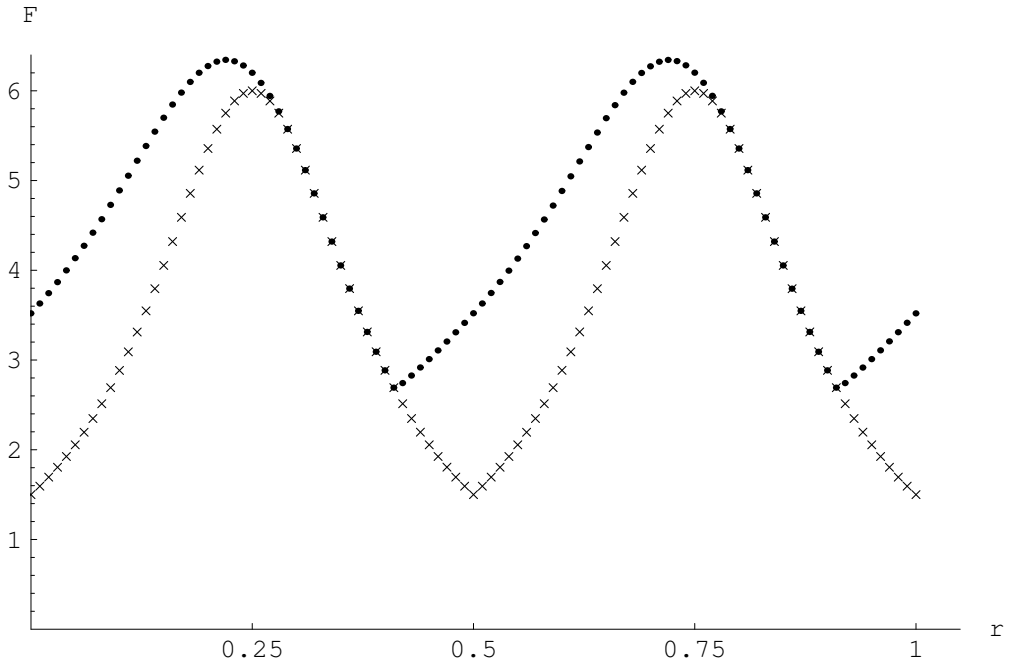


Figure 2: The merit factor of the r-rotated Legendre sequence of length 259499 before ($\times$) and after ($\bullet$) appending of the optimal number of its own initial elements, for varying r.

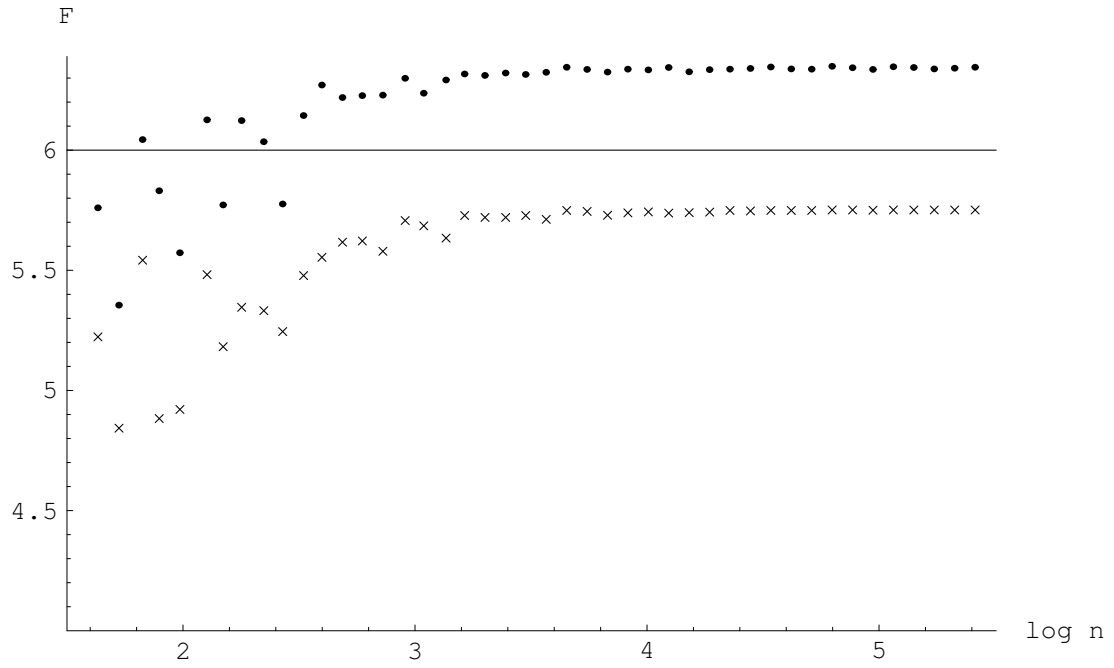


Figure 3: The merit factor of the (0.22)-rotated Legendre sequence of length n before ($\times$) and after ($^\bullet$) appending of the optimal number of its own initial elements, for varying n.

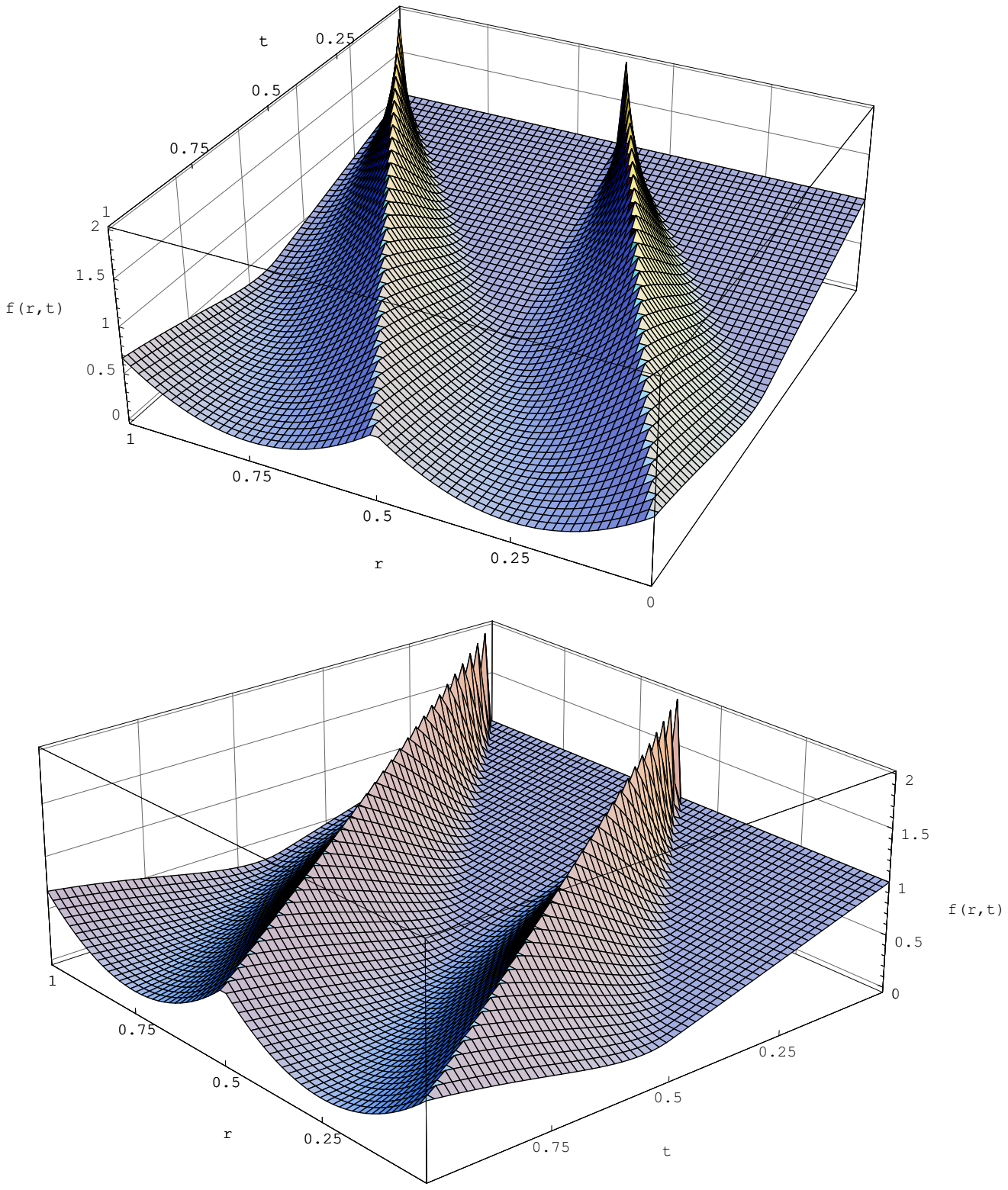


Figure 4: Variation of $f(r,t)$ with r and t from two viewpoints, according to Working Assumption 7.4.

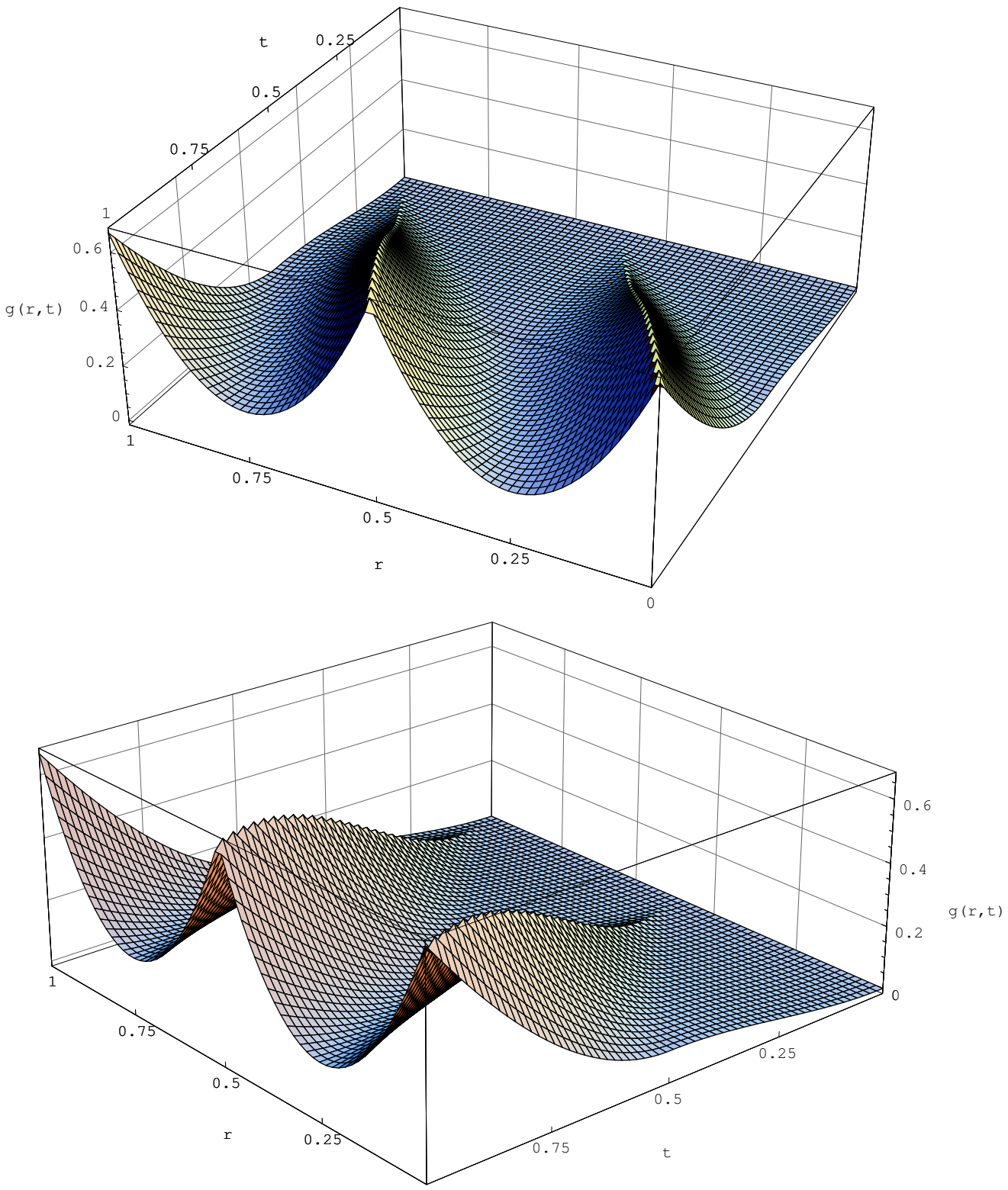
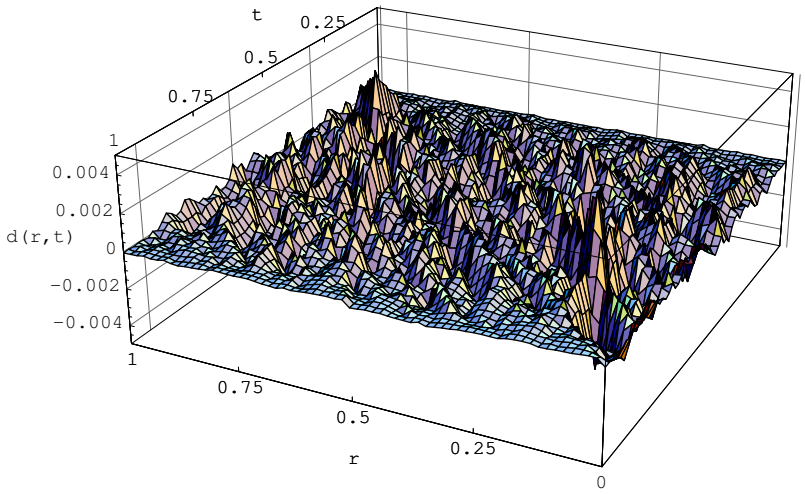
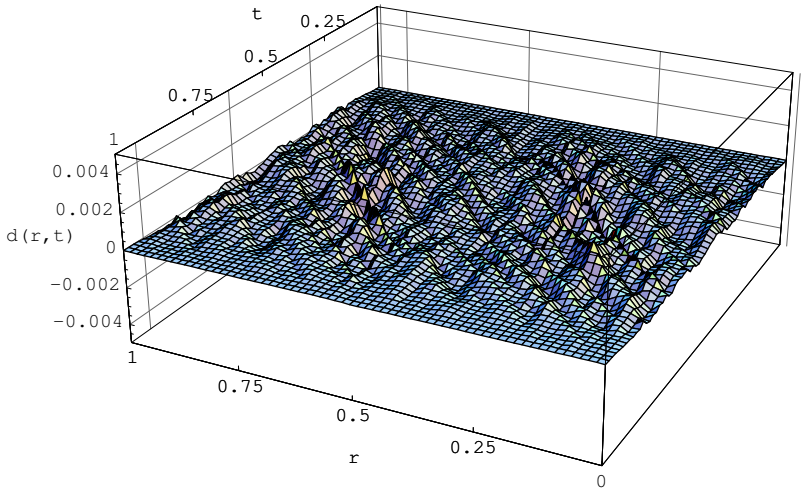


Figure 5: Variation of $g(r,t)$ with r and t from two viewpoints, according to Conjecture 7.5.

$n = 22783$



$n = 259499$



$n = 4433701$

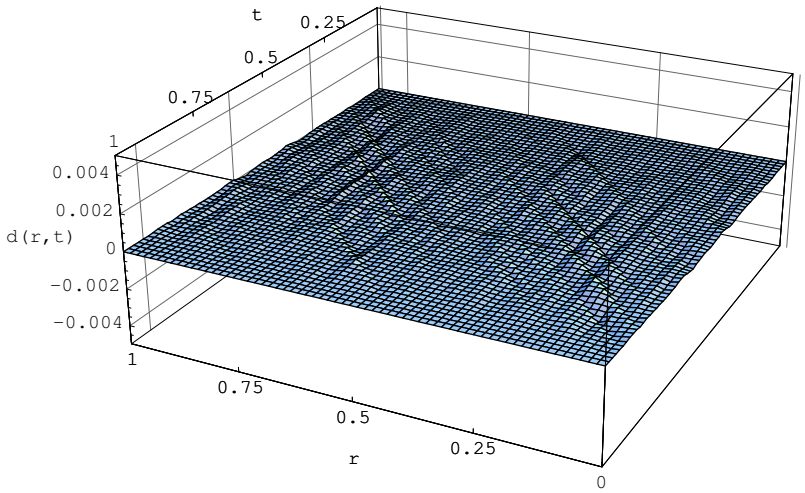


Figure 6: Discrepancy $d(r,t)$ between data value $t^2/F((X_r)^t)$ and $g(r,t)$ given by Conjecture 7.5, r and t varying in increments of $\frac{1}{64}$ and X a Legendre sequence of length n.

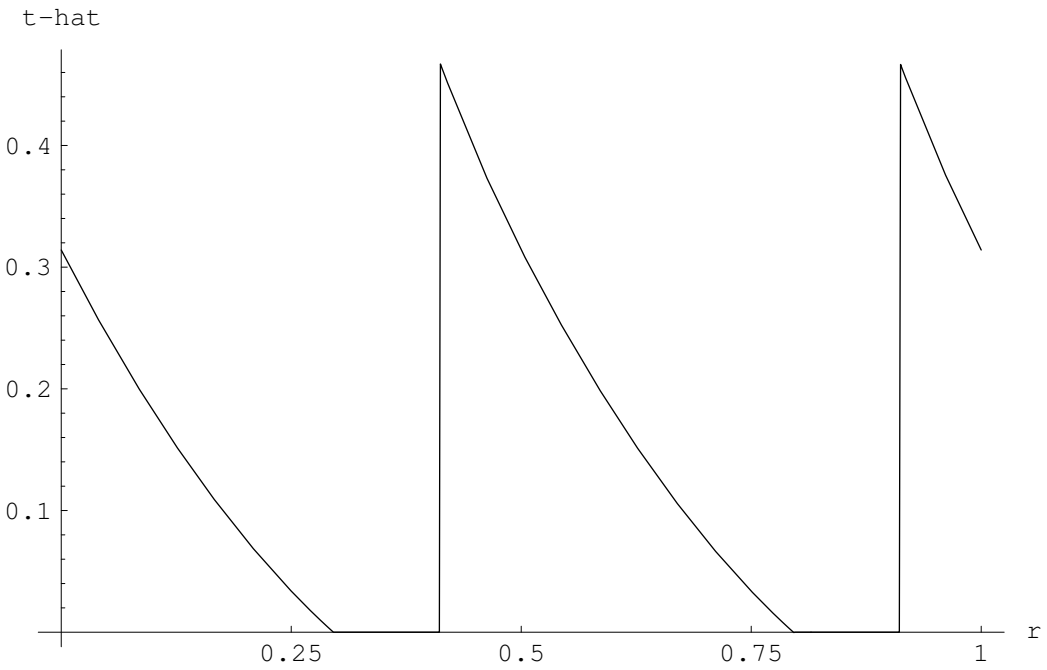


Figure 7: Variation with r of optimal value $\hat{t}$ maximizing $\lim_{n\longrightarrow\infty} F(X_r;(X_r)^t)$, according to Conjecture 7.5.

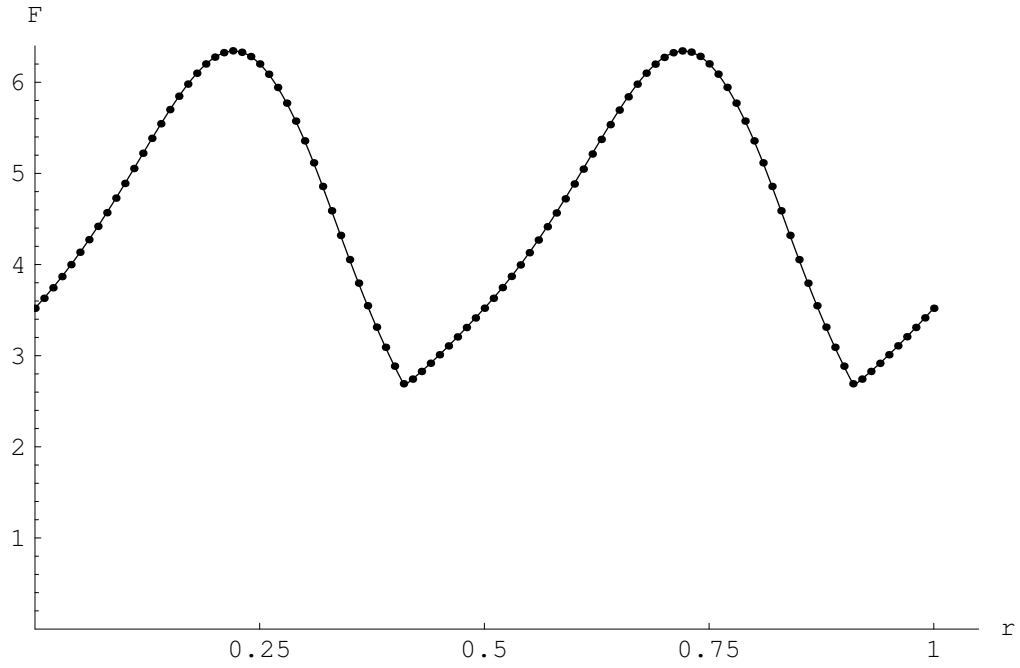


Figure 8: Comparison of $\lim_{n\longrightarrow\infty} F(X_r;(X_r)^{\hat{t}})$ according to Conjecture 7.5 (solid line) with $\max_{t\in[0,1]} F(X_r;(X_r)^t)$ at length $n = 259499$ (data points), as r varies.