18 February 1991

Final Report
ESTEC Contract No. 8696/89/NL/US   Work Order No.1
Technical Assistance for the CDMA Communication System Analysis

Period Reported: 1 February 1990 - 31 July 1990

WP 2110 **Investigation on an Optimal Code Family to be used in Synchronous High Capacity CDMA Communication Systems**
WP 2210 **Relation between Spreading Sequences and Error-Correcting Codes**

J.L. Massey, T. Mittelholzer
Institute for Signal- and Information Processing
CH-8092 ETH - Zürich

Contents

1 Introduction

From the point of view of code sequence design, one of the main novelties of the code-division multiple-access (CDMA) mobile communication system proposed by ESTEC is the synchronization of the users. In the forward link of this satellite communications system, all user sequences will arrive, of course, with no mutual time delays at the receiver end and no synchronization is necessary. In the return link from the mobile users to the satellite, ESTEC proposes to synchronize the users such that all sequences are almost aligned at the satellite (cf. [1]). This study considers how to select an optimal set of code sequences for such a quasi-synchronous (QS-) CDMA system on the return link in which the mutual time delays between sequences are less than **one** chip. Hereafter, only the return-link system is considered.

In the next section, basic notions such as the code sequence set are introduced. The principal aim of this section is to give an explicit model of **ordinary CDMA** and to describe the discrete channel seen by each user. It is important to note that the way that one processes the received signal strongly determines the capacity of the system. The matched-filter processing of each user in ordinary CDMA, which is the processing proposed in the current ESTEC CDMA application, is simple but in general not optimal. However, at present, it is an open question how to increase system capacity by using more efficient **practical** decoding methods.

The performance of ordinary CDMA depends strongly on the code sequence set that is used to spread the data. The performance can be expressed in terms of the maximal allowed time offset and of the crosscorrelation functions of the code sequences. Three optimality criteria for code selection are derived that differ by the synchronization assumption made for the system. All three conditions on the code set can be stated solely in terms of the even crosscorrelation functions of the code sequences because the system is almost synchronous. No odd crosscorrelation functions need to be considered, which greatly simplifies the search for optimal codes.

The basic limitation on the desired good even crosscorrelation properties comes from the Welch bound, which is a lower bound on the maximal crosscorrelation value of a code sequence set. The basic form of this important bound is derived herein in a simple way and, furthermore, a new condition for having equality in the Welch bound is presented. Starting from the basic form of the Welch bound, a new bound suited for QS-CDMA code sequences is derived. It is this bound that limits the maximal number of users of a QS-CDMA system when the code sequences have to have particularly good crosscorrelation properties around the origin, as is required in the current ESTEC CDMA application. Furthermore, the Welch bound implies that an A-CDMA system cannot fulfill the ESTEC requirements on the smallness of the crosscorrelation functions.

In Section 2.4, the preferentially-phased Gold code set is defined and investigated for its suitability in the current ESTEC QS-CDMA application. It is shown that the preferentially-phased Gold code set is optimal for S-CDMA, and essentially optimal for QS-CDMA, with respect to the formerly derived optimality criteria. For the preferentially-phased Gold code set, the crucial performance parameter of the proposed QS-CDMA system that governs the interuser interference is explicitly determined. It is important to note that this performance parameter strongly depends on whether the QS-CDMA system has slowly-varying or rapidly-varying phase offsets. The rapidly-varying case is advantageous, yielding an interference parameter that is about 6 times (7.8 dB) smaller than in the slowly-varying case.

In Section 3, the relation between error-correcting coding and spectrum spreading is considered. It is shown for the additive white Gaussian noise (AWGN) channel that the asymptotic coding gain of the coding system is independent of the spreading factor, both in the conventional CDMA systems as well as in unconventional CDMA systems of the QUALCOMM type. The relation among uniform convolutional codes, orthogonal convolutional codes, superorthogonal convolutional codes and equidistant partial-unit-memory (EPUM) convolutional codes is specified. The last of these classes is shown to provide the greatest asymptotic coding gain for a fixed decoder complexity. Finally, it is shown that the QUALCOMM unconventional spreading scheme is essentially optimum for many users. However, when the number of users is on the order of the spreading factor L and when either S-CDMA or QS-CDMA is used, it is shown that conventional CDMA using the preferentially-phased Gold sequence set is essentially optimum and substantially outperforms the unconventional CDMA system.

2 An Optimal Code Family for Quasi-Synchronous-CDMA

2.1 General Assumptions and Definitions

The code sequences (or spread-spectrum sequences) for use in the ESTEC CDMA application are two-valued. A **code sequence** of length L will be denoted by

$$\mathbf{x} = [x_0, x_1, ..., x_{L-1}] \qquad \text{where } x_i = \pm 1 \text{ for all } i.$$

The set of all code sequences, say M sequences, is the **code sequence set**

$$\mathcal{C} = \{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, ..., \mathbf{x}^{(M)}\}.$$

T will denote the **cyclic (left) shift operator**; it acts on a sequence of length L by the rule $T[x_0, x_1, ..., x_{L-1}] = [x_1, x_2, ..., x_{L-1}, x_0]$. The **even crosscorrelation function** $C_{\mathbf{xy}}(k)$ between two code sequences $\mathbf{x}, \mathbf{y} \in \mathcal{C}$ is defined by

$$C_{\mathbf{xy}}(k) =< \mathbf{x}, T^k \mathbf{y} >$$

where k is an integer and $< \cdot, \cdot >$ denotes the standard scalar product in the Euclidean space $\mathbf{R}^L$, i.e., $< \mathbf{x}, \mathbf{y} >= \sum_{i=0}^{L-1} x_i y_i$. The even crosscorrelation function $C_{\mathbf{xy}}(k)$ is periodic with period L and satisfies the following symmetry property:

$$C_{\mathbf{yx}}(k) = C_{\mathbf{xy}}(L-k) \text{ for all } k.$$

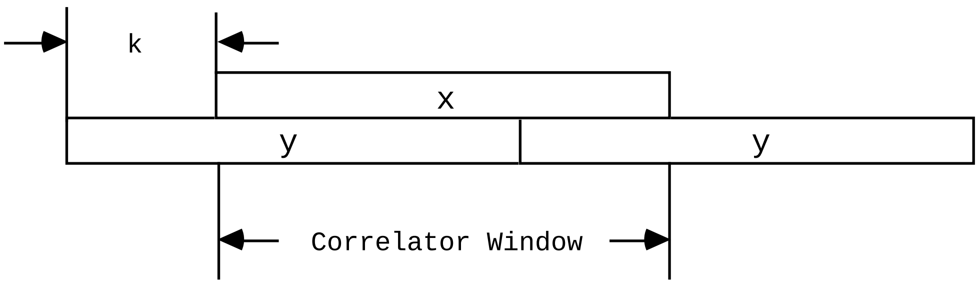

Fig. 2.1.1: Interpretation of $C_{\mathbf{xy}}(k)$.

In a similar way, using the **negacyclic (left) shift operator** N specified by $N[x_0, x_1, ..., x_{L-1}] = [x_1, x_2, ..., x_{L-1}, -\boldsymbol{x}_0]$, one can define the **odd crosscorrelation function** $\Theta_{\mathbf{xy}}(k)$ by

$$\Theta_{\mathbf{xy}}(k) =< \mathbf{x}, N^k \mathbf{y} > .$$

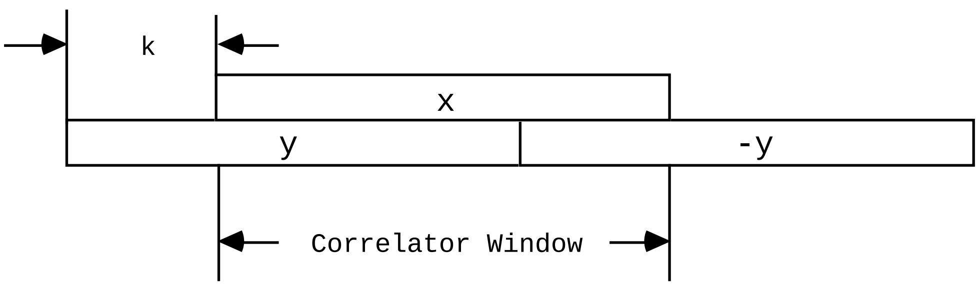

Fig. 2.1.2: Interpretation of $\Theta_{\mathbf{xy}}(k)$.

In general, the even and the odd crosscorrelation functions take on very different values for arbitrary k, $0 \leq k < L$. However, for $k = 1$, one has $\mathcal{C}_{\mathbf{xy}}(1) - \Theta_{\mathbf{xy}}(1) = 2x_{L-1}y_0 = \pm 2$ and hence

$$\mathcal{C}_{\mathbf{xy}}(1) \approx \Theta_{\mathbf{xy}}(1).$$

Thus, if one is interested in crosscorrelation properties of code sequences around the origin, say $k \in \{0, 1\}$, it is enough to consider only even crosscorrelation functions. Note that there are many analytical results for the even crosscorrelation functions, but only a few for the more complicated odd crosscorrelation functions. This simple but very useful fact is one of the important advantages from the point of view of CDMA sequence design that accrues from choosing a CDMA system where all the code sequences are (almost) aligned.

In the current ESTEC CDMA application, each "actual user" operates a direct-sequence/quadrature-phase-shift-keying (DS/QPSK) modulator, but uses the in-phase (I) and quadrature (Q) channels separately to transmit two independent **binary** data sequences, each with its own spreading sequence. Thus, with no loss of essential generality, we can consider that each "actual user" corresponds to two "virtual users", each of which has only one channel (I or Q), one data sequence, and one spreading sequence. Hereafter, the term **user** will mean one such "virtual user".

The data symbols $..., b_{-1}, b_0, b_1, b_2, ..., (b_i = \pm 1)$ of the i-th user are spread with the i-th code sequence, yielding an infinite chip stream

$$..., b_{-1} \cdot \mathbf{x}^{(i)}, b_0 \cdot \mathbf{x}^{(i)}, b_1 \cdot \mathbf{x}^{(i)}, b_2 \cdot \mathbf{x}^{(i)}, ... \qquad (1)$$

where $b_j \cdot \mathbf{x}^{(i)} = [b_j x_0^{(i)}, b_j x_1^{(i)}, ..., b_j x_{L-1}^{(i)}]$. Hence, the data symbol sequence acts as a polarity sequence on the code sequence. The **data symbol duration** T_s is given by

$$T_s = L \cdot T_c$$

where T_c denotes the **chip duration**. We assume hereafter that the receiver **symbol energy** has same value, E_s, for all users.

The performance of an ordinary CDMA system depends on whether the code sequences all arrive aligned or not at the receiver end. The **chip offsets (or time offsets)** of the sequences are measured by reference to a hypothetical **system clock** that provides the system with an absolute time reference. In general, the chip offset δ_i of a sequence $\mathbf{x}^{(i)} \in \mathcal{C}$ is a real number $0 \leq \delta_i < L$.

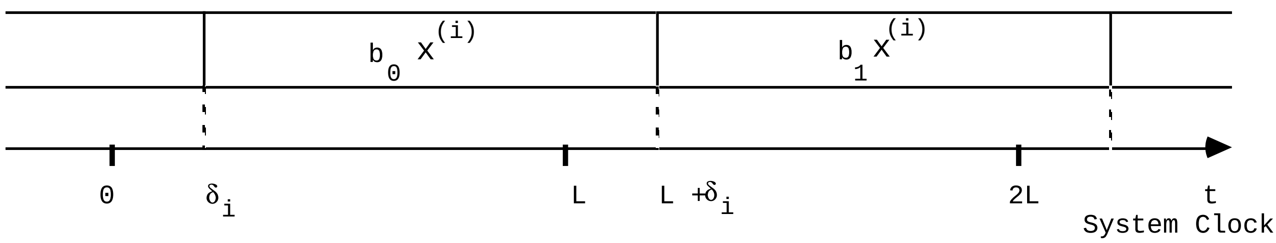

Fig. 2.1.3: Chip offset δ_i of sequence $\mathbf{x}^{(i)}$ with Respect to System Clock.

If all sequences are aligned with the system clock, i.e., if $\boldsymbol{\delta} \stackrel{def.}{=} [\delta_1, \delta_2, ..., \delta_M] = [0, 0, ..., 0]$, the system is called a **Synchronous-CDMA (S-CDMA)** system.

A **Quasi-Synchronous-CDMA (QS-CDMA)** system is characterized by the condition that the chip offsets for user i satisfies $0 \leq \delta_i \leq \Delta$ for all i and some $\Delta \leq 1$. The **relative chip offset** δ_{ij} between user i and user j is given by $\delta_{ij} = \delta_j - \delta_i$. Hence, one has

$$-\Delta \leq \delta_{ij} \leq \Delta.$$

In an ordinary CDMA system, the receiver for user i will apply a phase-coherent matched filter for user i to the received signal. The matched filter is assumed to be perfectly chip and carrier-phase synchronized with the sequence $\mathbf{x}$ of user i. Let ϕ_i denote the **carrier-phase offset** of user i with respect to the hypothetical system clock. When this matched filter for user i, who uses the code sequence $\mathbf{x}$, is applied over the time interval $\delta_i \leq t < T_s + \delta_i$ to the chip stream of user j, who uses the code sequence $\mathbf{y}$ and has a relative chip offset $\delta_{ij} = \delta_j - \delta_i$ and a relative carrier-phase offset $\phi_{ij} \stackrel{def.}{=} \phi_j - \phi_i$, then the following interfering output $(C_{ij})_0$ will result:

If the consecutive data symbols of user j are the same, say $b_0 = b_1 = 1$, and $\delta_{ij} \leq 0$, then

$$\begin{aligned} (C_{ij})_0 &= [|\delta_{ij}|(x_{L-1}y_0 + x_0 y_1 + ... + x_{L-2}y_{L-1}) + \\ &\quad (1 - |\delta_{ij}|)(x_0 y_0 + x_1 y_1 + ... + x_{L-1}y_{L-1})]cos(\omega\delta_{ij}T_c + \phi_{ij}) \\ (C_{ij})_0 &= [|\delta_{ij}|C_{\mathbf{xy}}(1) + (1 - |\delta_{ij}|)C_{\mathbf{xy}}(0)]cos(\omega\delta_{ij}T_c + \phi_{ij}), \end{aligned} \tag{2}$$

where the factor $cos(\omega\delta_{ij}T_c + \phi_{ij})$ results from the integration

$$\int_{\delta_i T_c}^{T_s+\delta_i T_c} cos(\omega(t-\delta_i T_c)-\phi_i)\cdot cos(\omega(t-\delta_j T_c)-\phi_j))dt \approx \frac{T_s}{2}cos(\omega(\delta_j-\delta_i)T_c+\phi_j-\phi_i) = \frac{T_s}{2}cos(\omega\delta_{ij}T_c+\phi_{ij}).$$

[This approximation is a virtual equality since the carrier-frequency ω satisfies $\omega \gg 1/T_c$.] If the consecutive data symbols are different, say $b_0 = 1$ and $b_1 = -1$, and $\delta_{ij} \leq 0$, then

$$(C_{ij})_0 = [|\delta_{ij}|\Theta_{\mathbf{xy}}(1) + (1 - |\delta_{ij}|)C_{\mathbf{xy}}(0)]cos(\omega\delta_{ij}T_c + \phi_{ij}). \tag{3}$$

If the chip offset δ_{ij} is positive, one gets in a similar way

$$(C_{ij})_0 = \begin{cases} [\delta_{ij}C_{\mathbf{xy}}(-1) + (1 - \delta_{ij})C_{\mathbf{xy}}(0)]\ cos(\omega\delta_{ij}T_c + \phi_{ij}) & \text{if } b_{-1} = b_0 = 1 \\ [\delta_{ij}\Theta_{\mathbf{xy}}(-1) + (1 - \delta_{ij})C_{\mathbf{xy}}(0)]\ cos(\omega\delta_{ij}T_c + \phi_{ij}) & \text{if } b_{-1} = -1, b_0 = 1. \end{cases} \tag{4}$$

Thus, the interfering outputs of the matched filters in a QS-CDMA system considered as functions of the code sequences depend only on the crosscorrelation functions $C_{\mathbf{xy}}(0)$ and $C_{\mathbf{xy}}(\pm 1)$ and the odd crosscorrelation functions $\Theta_{\mathbf{xy}}(0)$ and $\Theta_{\mathbf{xy}}(\pm 1)$. For QS-CDMA, it is enough to study the even crosscorrelation functions $C_{\mathbf{xy}}(k)$, $\mathbf{x}, \mathbf{y} \in \mathcal{C}$, around the origin, viz. $k \in \{0, \pm 1\}$ because $\Theta_{\mathbf{xy}}(\pm 1) \approx C_{\mathbf{xy}}(\pm 1)$ as was shown above.

The output of the matched filter for user i (normalized to have a unit-energy impulse response) due to his user i's signal is of course user i's amplitude times his ± 1 data symbol, i.e., $\sqrt{E_s}b_0^{(i)}$. The output due to the assumed additive white Gaussian noise on the channel is a Gaussian random variable, n_0, of mean 0 and variance $N_0/2$ where $N_0/2$ is the **two-sided noise power spectral density**. Thus, the total output of the matched filter for user i over one data symbol can be written

$$r_0^{(i)} = \sqrt{E_s}b_0^{(i)} + n_0 + \frac{\sqrt{E_s}}{L}c_0^{(i)}$$

where $c_0^{(i)} = \sum_{j=1,j\neq i}^{M}(C_{ij})_0$. In general, at discrete time k, the total output of the matched filter for user i will be

$$r_k^{(i)} = \sqrt{E_s}b_k^{(i)} + n_k + \frac{\sqrt{E_s}}{L}c_k^{(i)}$$

where the discrete-time sequence $\{n_k\}$ is a sequence of independent and identically-distributed (i.i.d.) Gaussian random variables having mean 0 and variance $N_0/2$.

It remains to consider the nature of the discrete-time interference sequences $\{c_k^{(i)}\}$ at the matched filter for user i. Note that $\frac{\sqrt{E_s}}{L}c_k^{(i)}$, which is called the **interuser interference** experienced by user i at discrete time k, depends according to (2) - (4) on the code sequence set $\mathcal{C}$, the carrier-phase offsets ϕ_i, the chip offsets δ_i and the data symbols $b_k^{(i)}$.

We shall assume that the relative carrier-phase offsets $\phi_{ij} = \phi_j - \phi_i$ and the relative chip offsets $\delta_{ij} = \delta_j - \delta_i$ are substantially constant over one symbol duration and that the corresponding discrete-time processes $\{(\phi_{ij})_k\}$ and $\{(\delta_{ij})_k\}$ are statistically independent. Furthermore, we assume that $\{(\phi_{ij})_k\}$ is an i.i.d. process with uniform distribution over the interval from 0 to 2π. Note that these assumtions are reasonable except for the two users corresponding to the "virtual user" pair that comes from the same "actual user". From these assumptions it follows that the discrete-time process $\{(Z_{ij})_k\}$, where $(Z_{ij})_k \stackrel{def.}{=} cos(\omega(\delta_{ij})_k T_c + (\phi_{ij})_k)$, is i.i.d. with mean 0 and variance 1/2. Moreover, the random variables $(Z_{ij})_k$ and $(\delta_{ij})_k$ are statistically independent (see Appendix B for a proof) and thus the random processes $\{(Z_{ij})_k\}$ and $\{(\delta_{ij})_k\}$ are also independent.

We assume that the discrete-time data sequence $\{b_k^{(i)}\}$ of each user is a coin-tossing sequence (thus also i.i.d.) of ± 1 valued random variables, and these sequences are indepenent for the different users. In terms of these discrete-time random variables, the interuser interference factor $c_k^{(i)}$ can be written as

$$\begin{aligned} c_k^{(i)} &= b_k^{(i)}\{ \sum_{\substack{j=1 \\ j\neq i,\delta_{ij}\leq 0}}^{M} b_k^{(j)}(|(\delta_{ij})_k| C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(1) + (1-|(\delta_{ij})_k|)C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(0))(Z_{ij})_k + \\ &\quad \sum_{\substack{j=1 \\ j\neq i,\delta_{ij}> 0}}^{M} b_k^{(j)}(|(\delta_{ij})_k| C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(-1) + (1-|(\delta_{ij})_k|)C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(0))(Z_{ij})_k\}. \end{aligned} \tag{5}$$

Note that, except for the relative chip offsets $(\delta_{ij})_k$, all discrete-time random variables occurring in the interuser interference are i.i.d.

We shall now distinguish two cases of QS-CDMA systems: A system having slowly-varying chip offsets $\{(\delta_i)_k\}$ and systems with rapidly-varying chip offsets. By slowly-varying chip offsets, we mean that the chip offset sequences $\{(\delta_i)_k\}$ are essentially constant. Thus, the relative chip offsets $\{(\delta_{ij})_k\}$ can therefore be treated as a stationary random process. Hence, for systems with slowly-varying chip offsets, the interuser interference process $\{\frac{\sqrt{E_s}}{L}c_k^{(i)}\}$ is definitely not i.i.d., but is constant in time.

A system with rapidly-varying chip offsets is characterized by having chip offsets $(\delta_i)_k$ that are independent from symbol to symbol and uniformly distributed in the interval from 0 to Δ. Thus, the relative chip offsets $\{(\delta_{ij})_k\}$ are i.i.d in this case and, therefore, the interuser interference $\frac{\sqrt{E_s}}{L}c_k^{(i)}$ is also i.i.d. for systems with rapidly-varying chip offsets. Note that in both cases we have a stationary system. In other words, the discrete-time channel of user i (compare Fig. 2.1.4) with binary antipodal input $\sqrt{E_s} b_k^{(i)}$, real-valued output $r_k^{(i)}$ and additive noise given by n_k and $(\sqrt{E_s}/L) \cdot c_k^{(i)}$ is memoryless for rapidly-varying chip-offsets. Because, for slowly-varying chip offsets, the chip offsets can be assumed to be constant in time, the discrete-time channel of user i is still memoryless, but the particular memoryless channel depends on the values assumed by these constant chip offsets.

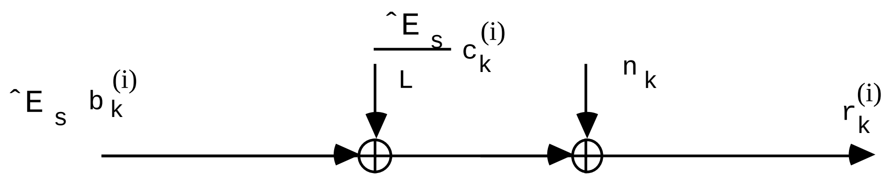

Fig. 2.1.4: Memoryless Discrete-Time Binary-Input/Real-Valued-Output Channel for User i.

A CDMA system, in which each user transmits over such a binary-input/continous-output channel as described above will be called an **ordinary CDMA** system. Note that processing the received signal independently for each user by matched filtering as specified in the channel model will in general entail a substantial loss of capacity compared to joint processing. A better use of the underlying channel could be made by applying joint processing at the receiving end, but little is presently known about such joint receivers that are simple enough to be practical.

In general, it is a complicated task to describe the probability distribution of the interuser interference factor $c_k^{(i)}$ in the above channel model. However, if there are many users and each contributes roughly the same amount to the interuser interference, then the central limit theorem asserts that the distribution of this random variable will be almost Gaussian. In case of slowly-varying chip offsets, the validity of the Gaussian approximation is easy to verify and it is a standard assumption in ordinary CDMA systems (cf. Chap. 1, Vol. II of [7]). For systems with rapidly-varying chip offsets, we shall justify the Gaussian approximation in Sec. 2.2.2 below; in fact the Gaussian assumption is even better in this case. Hence, for the comparision of different code sequence sets $\mathcal{C}$, the Gaussian assumption will be made in this report, i.e., it will be assumed that the random variables $c_k^{(i)}, i = 1, 2, ..., M$ are Gaussian. Under this assumption, the noise terms $(\sqrt{E_s}/L) \cdot c_k^{(i)}$ and n_k are independent Gaussian random variables with zero mean. Therefore, the total noise is again Gaussian with zero mean and variance equal to the sum of the variances of $(\sqrt{E_s}/L) \cdot c_k^{(i)}$ and n_k.

It is important to note that this Gaussian assumption for the interuser interference is conservative for **computing capacity** of the channel created by the matched-filter receiver because Gaussian noise is the worst type of additive noise for a given variance, i.e., it has the largest entropy.

An **Asynchronous-CDMA (A-CDMA)** system is characterized by allowing the chip offsets to be arbitrary; more precisely, the chip offsets δ_i are independent and uniformly distributed over the interval from 0 to L. In this case, even for rapidly-varying chip offsets, the equivalent discrete-time channel for bit decisions has a memory of one symbol, contrary to the memoryless discrete-time channel for S-CMDA or QS-CDMA. Thus, the discrete-time channel for A-CDMA is of an entirely different nature than the corresponding discrete-time channel for S-CDMA or QS-CDMA.

2.2 Criteria of Goodness

A set of code sequences $\mathcal{C}$ will be evaluated by considering the worst-user performance of an ordinary QS-CDMA system using this code, i.e., by the worst-user performance of the M independent binary-input/real-valued-output channels depicted in Fig. 2.1.4 when the Gaussian assumption is made for the interuser interference. Note that, *for the comparision of codes* $\mathcal{C}$, the thermal noise can be assumed to be zero, i.e., $n = 0$; thus, the performance depends only on the interuser interference factor $c^{(i)}, i = 1, 2, ..., M$ (here and in the sequel we drop the time subscripts, because the thermal noise and the interuser interference are both stationary processes). We shall treat the two cases of QS-CDMA systems (systems with slowly-varying chip offsets and systems with rapidly-varying chip offsets) separately.

2.2.1 Systems with Slowly-Varying Chip Offsets

A practical measure for the performance of an ordinary QS-CDMA system with *slowly-varying chip offsets* is the worst-user worst-case bit error probability

$$P_{wc} = \max_i \max_{\boldsymbol{\delta}} P_e(i),$$

where $P_e(i)$ is the bit error probability for uncoded random data transmission by user i and where the chip offsets may lie in the interval $0 \leq \delta_i \leq \Delta, i = 1, 2, ..., M,$ and $\Delta \leq 1$. The crucial aspect of a system with slowly-varying chip offsets is that the chip offset vector $\boldsymbol{\delta}$ can assume a worst-case value and remain in this condition over a substantial number of symbols so that the "worst-case" rather than the "average" interuser interference determines system performance. The worst-user worst-case error probability is determined by the largest variance of the interuser interference, i.e., by

$$\sigma_{wc}^2 = \max_i \max_{\boldsymbol{\delta}} Var(c^{(i)}). \tag{6}$$

Note that the worst-case variance of the interuser interference at the output of the channel in Fig.2.1.4 is $(E_s/L^2) \cdot \sigma_{wc}^2$. A more fundamental measure for the performance of a QS-CDMA system is the worst-user worst-case capacity for the M users, i.e.,

$$C_{wc} = \min_i \min_{\boldsymbol{\delta}} C_i,$$

where C_i denotes the capacity of the channel seen by user i. In a CDMA system, the users transmit independent data sequences. Furthermore, ordinary CDMA implies that the data sequence for each user is a coin-tossing sequence. [The coin-tossing distribution for $b_k^{(j)}$ in fact maximizes the variance of $c_k^{(i)}$ in (5).] Thus, again, the worst-case channel is determined by the largest variance among the interuser interferences (6). An upper bound on this worst-user worst-case capacity is the capacity of an additive Gaussian noise channel with a constraint E_s on the average input energy and is given by

$$C_{wc} \leq \frac{1}{2} \cdot log_2 \left(1 + \frac{E_s}{N_0/2 + (E_s/L^2) \cdot \sigma_{wc}^2}\right) = \frac{1}{2} \cdot log_2 \left(1 + \frac{1}{\frac{1}{2\gamma} + \frac{\sigma_{wc}^2}{L^2}}\right) \text{ bit/use },$$

where $\gamma = \frac{E_s}{N_0}$ denotes the signal-to-noise ratio of the channel when only one user is active, and where $(E_s/L^2) \cdot \sigma_{wc}^2$ denotes the variance of the worst-case interference. This upper bound is quite tight if the Gaussian assumption on interuser interference is valid and if E_s is not larger than $N_0/2 + (E_s/L^2) \cdot \sigma_{wc}^2$ so that a binary antipodal input nearly achieves capacity on an additive Gaussian noise channel with an average-input-energy constraint.

Note that the interuser interference factor $c^{(i)}$ has zero mean because all users transmit random data. Upon recalling that $Z_{ij} = cos(\omega \delta_{ij} T_c + \phi_{ij})$ has variance $1/2$, the variance of $c^{(i)}$ is readily obtained from formula (5) as

$$
\begin{aligned}
Var(c^{(i)}) &= \frac{1}{2} \sum_{\substack{j=1 \\ j \neq i, \delta_{ij} \leq 0}}^{M} (|\delta_{ij}| C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(1) + (1 - |\delta_{ij}|) C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(0))^2 + \\
& \quad \frac{1}{2} \sum_{\substack{j=1 \\ j \neq i, \delta_{ij} > 0}}^{M} (|\delta_{ij}| C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(-1) + (1 - |\delta_{ij}|) C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(0))^2 .
\end{aligned}
$$

The right side is a function of the code $\mathcal{C}$ and the phase-offset vector $\boldsymbol{\delta}$; it represents the product of the variance due to the phase offsets and the variance, which we will denote by $\sigma_i^2(\boldsymbol{\delta})$, due to the chip offsets and the random data where

$$
\begin{aligned}
\sigma_i^2(\boldsymbol{\delta}) &= \sum_{\substack{j=1 \\ j \neq i, \delta_{ij} \leq 0}}^{M} (|\delta_{ij}| C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(1) + (1 - |\delta_{ij}|) C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(0))^2 + \\
& \quad \sum_{\substack{j=1 \\ j \neq i, \delta_{ij} > 0}}^{M} (|\delta_{ij}| C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(-1) + (1 - |\delta_{ij}|) C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(0))^2 . \qquad (7)
\end{aligned}
$$

The $\sigma_i^2(\boldsymbol{\delta})$'s determine the crucial parameter σ_{wc}^2 that will be used to evaluate a code $\mathcal{C}$, viz.,

$$\sigma_{wc}^2 = \max_i \max_{\boldsymbol{\delta}} \frac{1}{2} \sigma_i^2(\boldsymbol{\delta}). \qquad (8)$$

The smaller σ_{wc}^2, the better the code, either in the sense of smaller worst-user worst-case bit error probability or in the sense of larger worst-user worst-case capacity.

General Optimality Criterion for Code Selection with Slowly-Varying Phase Offsets:

For a fixed number M of users, choose the code $\mathcal{C}$ that minimizes σ_{wc}^2.

For the special case of S-CDMA, which can be considered the special case of QS-CDMA when all chip offsets are all 0, this general criterion can be translated immediately into a criterion about crosscorrelation functions.

Optimality Criterion 1 for Code Selection for S-CDMA:

For a fixed number M of users, choose the code $\mathcal{C}$ that minimizes

$$\sigma_{wc}^2(0) = \max_{\mathbf{x}} \sum_{\substack{\mathbf{y}\in\mathcal{C} \\ \mathbf{y}\neq\mathbf{x}}} C_{\mathbf{xy}}^2(0). \tag{9}$$

We now return to QS-CDMA and want to translate the General Optimality Criterion, which depends on the continous parameter Δ, into a condition that depends only on the crosscorrelation values $C_{\mathbf{xy}}(0)$ and $C_{\mathbf{xy}}(\pm 1)$.

The maximum of $\sigma_i^2(\boldsymbol{\delta})$ in (7) is attained if all the terms of the sum are maximized independently. Moreover, if $\sigma_i^2(\boldsymbol{\delta})$ is maximal, then for each $j = 1, 2, ..., M$, one has either $\delta_j = 0$ or $\delta_j = \Delta$. If the chip offset of user i equals 0, then $\delta_{ij} \geq 0$ and one can write

$$\begin{aligned}\sigma_i^2(\boldsymbol{\delta}) &= \sum_{j\neq i, \delta_{ij}\geq 0} (\delta_{ij} C_{\mathbf{x}^{(i)}\mathbf{x}^{(j)}}(-1) + (1+\delta_{ij}) C_{\mathbf{x}^{(i)}\mathbf{x}^{(j)}}(0))^2 \\ &= \sum_{j\neq i, \delta_j = \Delta} \Delta^2 C^2_{\mathbf{x}^{(i)}\mathbf{x}^{(j)}}(-1) + \sum_{j\neq i, \delta_j = 0} C^2_{\mathbf{x}^{(i)}\mathbf{x}^{(j)}}(0).\end{aligned}$$

If $\delta_i = \Delta$, one gets

$$\sigma_i^2(\boldsymbol{\delta}) = \sum_{j\neq i, \delta_j = 0} \Delta^2 C^2_{\mathbf{x}^{(i)}\mathbf{x}^{(j)}}(1) + \sum_{j\neq i, \delta_j = \Delta} C^2_{\mathbf{x}^{(i)}\mathbf{x}^{(j)}}(0).$$

Letting $\Delta = 1$, one obtains the following criterion:

Optimality Criterion 2 for Code Selection for QS-CDMA with Slowly-Varying Chip Offsets:

For a fixed number M of users, choose the code $\mathcal{C}$ that minimizes

$$\max_i \max_{\boldsymbol{k}} \sum_{j\neq i} C^2_{\mathbf{x}^{(i)}\mathbf{x}^{(j)}}(k_j) \tag{10}$$

where $\boldsymbol{k} = [k_1, k_2, ..., k_M] \in \{0,1\}^M \cup \{0,-1\}^M$, i.e., where $\boldsymbol{k}$ is either a vector of M components chosen from $\{0,1\}$ or a vector of M components chosen from $\{0,-1\}$.

Remark: Note that, for very small Δ, the crosscorrelation values at the origin determine the magnitude of $\sigma_i^2(\boldsymbol{\delta})$. Thus, when evaluating a code for QS-CDMA, Criterion 1 for code selection for S-CDMA should also be considered, i.e., *one should, among nearly equally-optimal codes by Criterion 2, choose that code that is best according to Criterion 1.*

2.2.2 Systems with Rapidly-Varying Chip Offsets

Recall that in an ordinary QS-CDMA system with rapidly-varying chip offsets, the chip offset vector $\boldsymbol{\delta}$ changes independently from symbol to symbol and the components δ_i are independent and uniformly distributed in the interval from 0 to Δ, where $0 < \Delta \leq 1$. Recall that, $Z_{ij} = cos(\omega \delta_{ij} T_c + \phi_{ij})$ is well modeled as a zero-mean random variable with variance $\frac{1}{2}$; moreover, this random variable Z_{ij} can be assumed to be statistically independent of δ_{ij}. Note that in case of rapidly-varying chip offsets, the stochastic nature of the interuser interference $c^{(i)}$ given in (5) differs from the slowly-varying case where the relative chip offsets δ_{ij} were assumed to be essentially constant. We have to show that the Gaussian assumption can still be made when the interference $c^{(i)}$ depends on the random variable $\boldsymbol{\delta}$.

Considering $c^{(i)}$ as depending only on the random variable $\boldsymbol{\delta}$, one sees from (5) that it is the sum of $M-1$ random variables, say D_{ij}, with means 0 and variances depending only on the crosscorrelation values $C_{\mathbf{X}^{(j)}\mathbf{X}^{(i)}}(0)$ and $C_{\mathbf{X}^{(j)}\mathbf{X}^{(i)}}(\pm 1), i \neq j$. Because the random variables $\delta_{ij}, j = 1, 2, ..., M, j \neq i$, and $Z_{ij} = cos(\omega \delta_{ij} T_c + \phi_{ij}), j = 1, 2, ..., M, j \neq i$, are independent, and because the δ_{ij}'s, resp. the Z_{ij}'s, are identically distributed, the D_{ij}'s are also statistically independent and "almost" identically distributed. Thus, the Gaussian approximation is justified for large M, i.e., the interuser interference can be assumed to be Gaussian when considered as a function of the random data symbols **and** the independent and uniformly distributed chip offsets **and** the independent and uniformly distributed carrier-phase offsets. Because the data symbols $\boldsymbol{b}$ and the chip offsets $\boldsymbol{\delta}$ and the random variables $Z_{ij} = cos(\omega \delta_{ij} T_c + \phi_{ij})$ are statistically independent, the variance of the interference is given by

$$Var(c^{(i)}) = E_{\boldsymbol{\delta}}[E_{\boldsymbol{b}}[(c^{(i)})^2]] = Var(cos(\omega \delta_{ij} T_c + \phi_{ij})) E_{\boldsymbol{\delta}}[\sigma_i^2(\boldsymbol{\delta})] = \frac{1}{2} E_{\boldsymbol{\delta}}[\sigma_i^2(\boldsymbol{\delta})],$$

where $\sigma_i^2(\boldsymbol{\delta})$ is still given by (7). Thus, the crucial parameter to measure the performance of a code is

$$\sigma^2 = \max_i \frac{1}{2} E_{\boldsymbol{\delta}}[\sigma_i^2(\boldsymbol{\delta})].$$

We shall now compute the average variance $\frac{1}{2}E_{\boldsymbol{\delta}}[\sigma_i^2(\boldsymbol{\delta})]$ seen by user i explicitly in terms of Δ and the crosscorrelation functions. Because the chip offsets are independent and uniformly distributed in the interval from 0 to Δ, it follows that the relative chip offsets $\delta_{ij}, i \neq j$, are also independent and have identical distributions with density functions

$$p_{\delta_{ij}}(\lambda) = \begin{cases} \frac{1}{\Delta^2}(\Delta - \lambda) & \text{for } 0 \leq \lambda \leq \Delta \\ \frac{1}{\Delta^2}(\Delta + \lambda) & \text{for } -\Delta \leq \lambda < 0. \end{cases}$$

The expectation of a typical term - from the interference of user i with user j - of $\sigma_i^2(\boldsymbol{\delta})$ can be computed as follows:

$$\begin{aligned} V_{ij} &= E\left[(|\delta_{ij}|C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(\pm 1) + (1 - |\delta_{ij}|)C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(0))^2\right] \\ &= E\left[(|\delta_{ij}|C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(1) + (1 - |\delta_{ij}|)C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(0))^2 \middle| \delta_{ij} \leq 0\right] P(\delta_{ij} \leq 0) \\ &\quad E\left[(|\delta_{ij}|C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(-1) + (1 - |\delta_{ij}|)C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(0))^2 \middle| \delta_{ij} > 0\right] P(\delta_{ij} > 0). \end{aligned}$$

The first term of the last sum is given by

$$\int_{-\Delta}^{0} (|\lambda|C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(1) + (1 - |\lambda|)C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(0))^2 p_\lambda(\lambda) d\lambda$$

$$= \frac{\Delta^2}{12} C^2_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(1) + (\frac{1}{2} - \frac{\Delta}{3} + \frac{\Delta^2}{12}) C^2_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(0) + (\frac{\Delta}{3} - \frac{\Delta^2}{6}) C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(1) C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(0).$$

The second term of the sum for V_{ij} is obtained by replacing $C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(1)$ by $C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(-1)$ in the equation above. The average variance seen by user i can now be written as

$$\begin{aligned} \frac{1}{2} E[\sigma_i^2(\boldsymbol{\delta})] &= \frac{1}{2} \sum_{j \neq i} V_{ij} \\ &= \frac{1}{2} \sum_{j \neq i} \{ \frac{\Delta^2}{12} (C^2_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(1) + C^2_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(-1)) + (1 - \frac{2\Delta}{3} + \frac{\Delta^2}{6}) C^2_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(0) + \\ & (\frac{\Delta}{3} - \frac{\Delta^2}{6}) C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(0) (C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(-1) + C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(1) \}. \end{aligned} \tag{11}$$

Note that this average variance is a sum of the quadratic forms V_{ij} in the variables $C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(-1)$, $C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(0)$ and $C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(1)$. Collecting the terms in the quadratic forms V_{ij} differently, i.e., diagonalizing the quadratic forms, one gets

$$\begin{aligned} \frac{1}{2} E[\sigma_i^2(\boldsymbol{\delta})] &= \frac{1}{6} \sum_{j \neq i} \{ ((\frac{\Delta}{2}) C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(1) + (1 - \frac{\Delta}{2}) C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(0))^2 + \\ & ((\frac{\Delta}{2}) C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(-1) + (1 - \frac{\Delta}{2}) C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(0))^2 + C^2_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(0) \}. \end{aligned} \tag{12}$$

Letting $\Delta = 0$, the optimality criterion for σ^2 reduces to Criterion 1 (for S-CDMA). Thus, for very small Δ, one should use the optimality condition as Criterion 1 for the code set.

Now consider the case $\Delta \approx 1$. The average variance (12) is still dominated by the crosscorrelation values at the origin. Thus, one should require that these values be small, i.e., the code should still be optimal with respect to Criterion 1. If these values $C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(0)$ are all very small, one can neglect them in the first and second terms on the right in equation (12) so that optimizing σ^2 can then be translated to the following condition on the code sequence set $\mathcal{C}$.

Optimality Criterion 3 for Code Selection for QS-CDMA with Rapidly-Varying chip offsets:

For a fixed number M of users, choose the code $\mathcal{C}$ that minimizes

$$\max_i \sum_{j \neq i} \{ C^2_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(-1) + C^2_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(0) + C^2_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(+1) \}. \tag{13}$$

Note that, for $\Delta = 2$, this optimality criterion would follow directly from (12) without any assumptions on the smallness of the crosscorrelation values $C_{\mathbf{X}^{(i)}\mathbf{X}^{(j)}}(0)$.

2.3 The Welch Bound and its Implications

The search for optimal code sets with respect to Criteria 1, 2 or 3 is a deterministic problem concerning the even crosscorrelation functions of the code sequences. The aim of the optimization is to find sequences having small crosscorrelation values around the origin. But, all these values cannot be made arbitrarily small in general because each ± 1 sequence satisfies the condition

$$C_{\mathbf{xx}}(0) = <\mathbf{x}, \mathbf{x}> = L,$$

where L denotes the length of the sequence $\mathbf{x}$. This constraining phenomen, which holds even when the sequences are not ± 1 valued but still all have the same energy, is a well-known fact (cf. [3], [4]) and the bound that is usually cited in connection with it is called the Welch bound; as usually cited, it lower bounds (see e.g. [3]) the maximal crosscorrelation value of a code $\mathcal{C}$ defined as

$$c_{max} = \max_{\mathbf{x} \in \mathcal{C}} \max_{\substack{\mathbf{x} \in \mathcal{C} \\ \mathbf{y} \neq \mathbf{x}}} \max_{k} |C_{\mathbf{xy}}(k)|.$$

As we are interested rather in a bound on the average crosscorrelation parameters σ_{wc}^2 or σ^2, we shall go back to the basic form of the Welch bound, which is actually a result about the magnitude of scalar products of a fixed number of real-valued vectors all having the same norm (cf. [3]). Starting from this basic form, some new convenient bounds for σ_{wc}^2 in the QS case can be derived. We commence with a new and simple derivation of this basic inequality that will also give insight into a new fundamental condition for the Welch bound to hold with equality for a $\pm$ 1 valued sequence set.

Lemma 1: For any real numbers $a_0, a_1, ..., a_{L-1}$,

$$\sum_{i=0}^{L-1} a_i^2 \geq \frac{1}{L} \left(\sum_{i=0}^{L-1} a_i \right)^2$$

with equality if and only if $a_0 = a_1 = ... = a_{L-1}$.

Proof. The Cauchy-Schwarz inequality $||\mathbf{a}||^2||\mathbf{b}||^2 \geq (<\mathbf{a}, \mathbf{b}>)^2$ with equality if and only if $\mathbf{a}$ and $\mathbf{b}$ are proportional, for $\mathbf{a} = [a_0, a_1, ..., a_{L-1}]$ and $\mathbf{b} = [\frac{1}{L}, \frac{1}{L}, ..., \frac{1}{L}]$ yields

$$\frac{1}{L} \cdot \sum_{i=0}^{L-1} a_i^2 \geq \left(\sum_{i=0}^{L-1} \frac{1}{L} a_i \right)^2$$

with equality if and only if $a_0 = a_1 = ... = a_{L-1}$.

Welch's Bound: If $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, ..., \mathbf{x}^{(M)}$ are vectors of energy E in $\mathbf{R}^L$ (i.e., $<\mathbf{x}^{(m)}, \mathbf{x}^{(m)}> = E$ for all m), then

$$\sum_{m=1}^{M} \sum_{n=1}^{M} <\mathbf{x}^{(m)}, \mathbf{x}^{(n)}>^2 \geq \frac{M^2}{L} E^2.$$

Proof.

$$\sum_{m=1}^{M} \sum_{n=1}^{M} <\mathbf{x}^{(m)}, \mathbf{x}^{(n)}>^2 = \sum_{m=1}^{M} \sum_{n=1}^{M} (\sum_{i=0}^{L-1} x_i^{(m)} x_i^{(n)})^2$$

$$\begin{aligned}
&= \sum_{m=1}^{M}\sum_{n=1}^{M}\sum_{i=0}^{L-1} x_i^{(m)} x_i^{(n)} \sum_{j=0}^{L-1} x_j^{(m)} x_j^{(n)} \\
&= \sum_{i=0}^{L-1}\sum_{j=0}^{L-1}\sum_{m=1}^{M} x_i^{(m)} x_j^{(m)} \sum_{n=1}^{M} x_i^{(n)} x_j^{(n)} \\
&= \sum_{i}\sum_{j}\left(\sum_{m} x_i^{(m)} x_j^{(m)}\right)^2 \\
&\overset{1)}{\geq} \sum_{i}\left(\sum_{m} (x_i^{(m)})^2\right)^2 \\
&\overset{2)}{\geq} \frac{1}{L}\left(\sum_{i}\sum_{m} (x_i^{(m)})^2\right)^2 \\
&= \frac{1}{L}\left(\sum_{m}\sum_{i} (x_i^{(m)})^2\right)^2 \\
&= \frac{1}{L}\left(\sum_{m} E\right)^2 = \frac{M^2E^2}{L}.
\end{aligned} \tag{14}$$

Inequality $^{1)}$ is obtained by keeping only the terms with $i = j$ and inequality $^{2)}$ follows from the lemma with $a_i = \sum_m (x_i^{(m)})^2$.

Remark: Note that when the M vectors are code sequences with components equal to ± 1 then $a_i = \sum_m (x_i^{(m)})^2 = M$ for all i and therefore inequality $^{2)}$ holds with equality. Thus, it is enough to consider inequality $^{1)}$ to decide under what condition the Welch bound holds with equality for a code sequence set $\mathcal{C}$.

Condition for Equality in the Welch Bound: The Welch bound holds with equality for a ± 1 code sequence set $\mathcal{C}$ if and only if the code sequences satisfy

$$\sum_{m=1}^{M} x_i^{(m)} x_j^{(m)} = 0 \;\text{ for all } i, j, i \neq j.$$

An illustrative interpretation of this condition is obtained, when one writes the code sequences as **rows** in an array

$$\begin{array}{cccc}
x_0^{(1)} & x_1^{(1)} & \cdots & x_{L-1}^{(1)} \\
x_0^{(2)} & x_1^{(2)} & \cdots & x_{L-1}^{(2)} \\
\vdots & \vdots & & \vdots \\
x_0^{(M)} & x_1^{(M)} & \cdots & x_{L-1}^{(M)}.
\end{array} \tag{15}$$

The condition now reads: The Welch bound holds with equality if and only if all **columns** of the array (15) are orthogonal.

Because of the normalizing condition $E =< \mathbf{x}, \mathbf{x} >= L$ and because $< \mathbf{x}^{(m)}, \mathbf{x}^{(n)} >^2 = C^2_{\mathbf{x}^{(m)}\mathbf{x}^{(n)}}(0)$, the basic form of the Welch bound translates to the following bound for a code sequence set:

$$\sum_{m=1}^{M}\sum_{n=1}^{M} C^2_{\mathbf{X}^{(m)}\mathbf{X}^{(n)}}(0) \geq M^2 L.$$

Collecting the terms $C^2_{\mathbf{X}^{(m)}\mathbf{X}^{(m)}}(0) = L^2$ for $m = 1, 2, ..., M$ separately, one gets

$$\sum_{m=1}^{M}\sum_{n\neq m} C^2_{\mathbf{X}^{(m)}\mathbf{X}^{(n)}}(0) \geq M^2 L - ML^2 = ML(M-L).$$

The inner sum on the left of the inequality can be upper bounded by M times the worst-case parameter $\sigma^2_{wc}(0)$ of (9) and the inequality above becomes

$$M\sigma^2_{wc}(0) \geq ML(M-L).$$

We have now derived the following result for S-CDMA codes, which gives a limitation on the smallness of the crucial parameter $\sigma^2_{wc}(0)$. Note that by Criterion 1, this is the parameter one wishes to minimize.

Welch Bound for S-CDMA Code Sequences:

In S-CDMA, the variance of the worst-user worst-case interuser interference is lower bounded by

$$\sigma^2_{wc}(0) \geq L(M-L). \tag{16}$$

Remark: For a "good" code, this bound on $\sigma^2_{wc}(0)$ is very tight, and for the proposed Gold code set it is in fact satisfied with equality (see Section 2.4, below).

As done above for the synchronous case, we now want to apply the Welch bound to the QS case. To this end, we consider the union of a code $\mathcal{C}$ and some of its cyclic shifts, viz.

$$\mathcal{C}_\Delta \stackrel{def.}{=} \mathcal{C} \cup T\mathcal{C} \cup ... \cup T^\Delta \mathcal{C},$$

where Δ denotes a specified **integer**, $0 \leq \Delta \leq L-1$. The set $\mathcal{C}_\Delta$ will be called the **virtual code set** associated to $\mathcal{C}$. Generalizing Criterion 2, which considers the special case $\Delta = 1$, we are interested in minimizing the worst-case variance

$$\sigma^2_{wc}(\Delta) \stackrel{def.}{=} \max_i \max_{\boldsymbol{k}} \sum_{j\neq i} < T^{k_i}\mathbf{x}^{(i)}, T^{k_j}\mathbf{x}^{(j)} >^2,$$

where the integer-valued phase-offset vector $\boldsymbol{k}$ lies in the cube $\{0, 1, ..., \Delta\}^M$. The maximal value $\Delta = L-1$ corresponds to A-CDMA when all users send a constant data sequence; recall, however, that if the users send random binary data, one also has to take the odd crosscorrelation functions into account unless Δ is small. In the special case of A-CDMA when all users send constant data, the worst-user worst-case interference parameter $\sigma^2_{wc}(L-1)$ depends only on the even correlation functions and thus will be a lower bound on the variance of the worst-user worst-case interference of an actual random-data A-CDMA system. If the code sequences have bad odd-crosscorrelation properties, this lower bound will not be tight.

The crucial parameter $\sigma^2_{wc}(\Delta)$ for system performance will be put in relation to the worst-user worst-case variance of the interuser interference of the virtual code set defined by

$$\sigma^2_\Delta \stackrel{def.}{=} \max_{\mathbf{x}\in\mathcal{C}_\Delta} \sum_{\substack{\mathbf{y}\in\mathcal{C}_\Delta \\ \mathbf{y}\neq\mathbf{x}}} < \mathbf{x},\mathbf{y} >^2 .$$

The virtual code set, consisting of $M(\Delta+1)$ code sequences, is assumed to be used in synchronous mode and therefore the Welch bound can be applied yielding

$$L((\Delta+1)M - L) \leq \sigma^2_\Delta. \tag{17}$$

It must be recalled, however, that only M of the $M\Delta$ code sequences in the virtual code set will actually be in use in the real QS-CDMA system. In the sum defining σ^2_Δ, we shall identify those terms that sum up to $\sigma^2_{wc}(\Delta)$ and, furthermore, we shall bound them in a way such as to obtain a useful relation between σ^2_Δ and $\sigma^2_{wc}(\Delta)$. Such a relation is useful if it allows one to lower bound $\sigma^2_{wc}(\Delta)$ using inequality (17).

By the definition of $\mathcal{C}_\Delta$ and σ^2_Δ, one can write

$$\sigma^2_\Delta = \max_i \max_l \left(\sum_{\substack{k=0\\k\neq l}}^{\Delta} < T^l\mathbf{x}^{(i)}, T^k\mathbf{x}^{(i)} >^2 + \sum_{j\neq i}\sum_{k=0}^{\Delta} < T^l\mathbf{x}^{(i)}, T^k\mathbf{x}^{(j)} >^2 \right).$$

Using the fact that a total mean can be expressed in terms of suitably weighted submeans, i.e.,

$$\frac{1}{A+B}\left(\sum_{m=1}^{A} a_m + \sum_{n=1}^{B} b_n\right) = \frac{A}{A+B}\frac{1}{A}\sum_{m=1}^{A} a_m + \frac{B}{A+B}\frac{1}{B}\sum_{n=1}^{B} b_n,$$

one obtains

$$\begin{aligned}
\frac{1}{(\Delta+1)M-1}\sigma^2_\Delta &= \max_i \max_l \left(\frac{\Delta}{(\Delta+1)M-1}\frac{1}{\Delta}\sum_{\substack{k=0\\k\neq l}}^{\Delta} < T^l\mathbf{x}^{(i)}, T^k\mathbf{x}^{(i)} >^2 + \right.\\
&\qquad \left. \frac{(\Delta+1)(M-1)}{(\Delta+1)M-1}\frac{1}{(\Delta+1)(M-1)}\sum_{j\neq i}\sum_{k=0}^{\Delta} < T^l\mathbf{x}^{(i)}, T^k\mathbf{x}^{(j)} >^2 \right)\\
&\leq \max_i \max_l \frac{\Delta}{(\Delta+1)M-1}\frac{1}{\Delta}\sum_{\substack{k=0\\k\neq l}}^{\Delta} < T^l\mathbf{x}^{(i)}, T^k\mathbf{x}^{(i)} >^2 + \\
&\qquad \max_i \max_l \frac{(\Delta+1)(M-1)}{(\Delta+1)M-1}\frac{1}{(\Delta+1)(M-1)}\sum_{j\neq i}\sum_{k=0}^{\Delta} < T^l\mathbf{x}^{(i)}, T^k\mathbf{x}^{(j)} >^2 .
\end{aligned}$$

Because the mean is smaller than the maximum, we have

$$\frac{1}{\Delta+1}\sum_{k=0}^{\Delta} < T^l\mathbf{x}^{(i)}, T^k\mathbf{x}^{(j)} >^2 \leq \max_{k_j} < T^{k_i}\mathbf{x}^{(i)}, T^{k_j}\mathbf{x}^{(j)} >^2 . \tag{18}$$

Thus, one obtains

$$\frac{1}{(\Delta+1)M-1}\sigma_\Delta^2 \leq \max_i \max_l \frac{\Delta}{(\Delta+1)M-1}\frac{1}{\Delta}\sum_{\substack{k=0\\k\neq l}}^{\Delta} < T^l\mathbf{x}^{(i)}, T^k\mathbf{x}^{(i)} >^2 + \max_i \max_{\boldsymbol{k}} \frac{(\Delta+1)(M-1)}{(\Delta+1)M-1}\frac{1}{M-1}\sum_{j\neq i} < T^{k_i}\mathbf{x}^{(i)}, T^{k_j}\mathbf{x}^{(j)} >^2 . \quad (19)$$

Defining the auxiliary parameter

$$\alpha^2 \overset{def.}{=} \max_i \max_l \sum_{\substack{k=0\\k\neq l}}^{\Delta} < T^l\mathbf{x}^{(i)}, T^k\mathbf{x}^{(i)} >^2,$$

which collects all autocorrelation terms, one can rewrite the last inequality in compact form as

$$\frac{\sigma_\Delta^2}{(\Delta+1)M-1} \leq \frac{\Delta}{(\Delta+1)M-1}\frac{\alpha^2}{\Delta} + \frac{(\Delta+1)(M-1)}{(\Delta+1)M-1}\frac{\sigma_{wc}^2(\Delta)}{M-1}. \quad (20)$$

For the interpretation of inequality (20) the following parameters are introduced:

$$\tilde{\sigma}_\Delta^2 \overset{def.}{=} \frac{\sigma_\Delta^2}{(\Delta+1)M-1}$$
$$\tilde{\sigma}_{wc}^2(\Delta) \overset{def.}{=} \frac{\sigma_{wc}^2(\Delta)}{M-1}$$
$$\tilde{\alpha}^2 \overset{def.}{=} \frac{\alpha^2}{\Delta}.$$

Note that these new parameters represent averaged auto- and crosscorrelation parameters on a **per user** basis, in particular, $\tilde{\sigma}_{wc}^2(\Delta) \leq c_{max}^2$, where c_{max} denotes the maximal crosscorrelation value of the code $\mathcal{C}$, and $\tilde{\alpha}^2$ is less or equal to the square of the maximal off-peak autocorrelation value of all code sequences. By (20), one can state a basic relation connecting the averaged per-user variances of a code sequence set $\mathcal{C}$ and its associated virtual code set $\mathcal{C}_\Delta$:

Basic Inequality: $\tilde{\sigma}_\Delta^2 \leq \frac{\Delta}{(\Delta+1)M-1}\tilde{\alpha}^2 + \frac{(\Delta+1)(M-1)}{(\Delta+1)M-1}\tilde{\sigma}_{wc}^2(\Delta).$

The Welch bound (17) for the synchronous virtual code set $\mathcal{C}_\Delta$ together with the Basic Inequality gives the following lower bound on $\tilde{\alpha}^2$ and $\tilde{\sigma}_{wc}^2(\Delta)$.

Extended Welch Bound for QS-CDMA Code Sequences:

In QS-CDMA, the averaged per-user auto- and crosscorrelation parameters $\tilde{\alpha}^2$ and $\tilde{\sigma}_{wc}^2(\Delta)$ obey the following bound:

$$\frac{L((\Delta+1)M-L)}{(\Delta+1)M-1} \leq \frac{\Delta}{(\Delta+1)M-1}\tilde{\alpha}^2 + \frac{(\Delta+1)(M-1)}{(\Delta+1)M-1}\tilde{\sigma}_{wc}^2(\Delta). \quad (21)$$

Remarks: 1. For the special case $\Delta = 0$, inequality (21) reduces to the Welch bound (16) for the synchronous case.

2. For $\Delta = L - 1$, this bound allows one to bound the **even** crosscorrelation functions in the asynchronous case.

3. If one replaces the cyclic shift operator by the negacyclic shift operator, the Basic Inequality still holds and one gets the same bound for the **odd** crosscorrelation functions as one has for the even ones. However, it might be possible to derive a larger (and thus better) lower bound on the maximum of both the even and odd correlation functions considered simultaneously; but this appears to be difficult.

4. Usually (cf. [5]) for the asynchronous case, the Welch bound is stated in terms of c_{max} and a_{max}, where c_{max} is the maximal crosscorrelation value (see above) and a_{max} denotes the largest offpeak magnitude of all autocorrelation functions:

$$\frac{L^2(M-1)}{LM-1} \leq max\{a_{max}^2, c_{max}^2\}.$$

Note that (21) implies the same bound for the averaged parameters $\tilde{\alpha}^2$ and $\tilde{\sigma}_{wc}^2(\Delta)$. Thus, the "Welch bound" is essentially a lower bound on the averaged parameters $\tilde{\alpha}^2$ and $\tilde{\sigma}_{wc}^2(\Delta)$ rather than on the larger values c_{max}^2 and a_{max}^2. It is well-known (cf. [5] or [6]) that the Welch bound is not tight for $max\{c_{max}^2, a_{max}^2\}$ but that a stronger inequality can be obtained for binary CDMA sequences. We can say little about the tightness of the new inequality (21), but it seems that it is also not very tight when $\Delta \neq 0$ (compare Table 2.5.2, below).

The nature of the Welch bounds (16) and (21) is best studied by considering the function

$$f(x) = \frac{L(x-L)}{x-1},$$

which appears in both bounds. Its behavior is illustrated in the following plot:

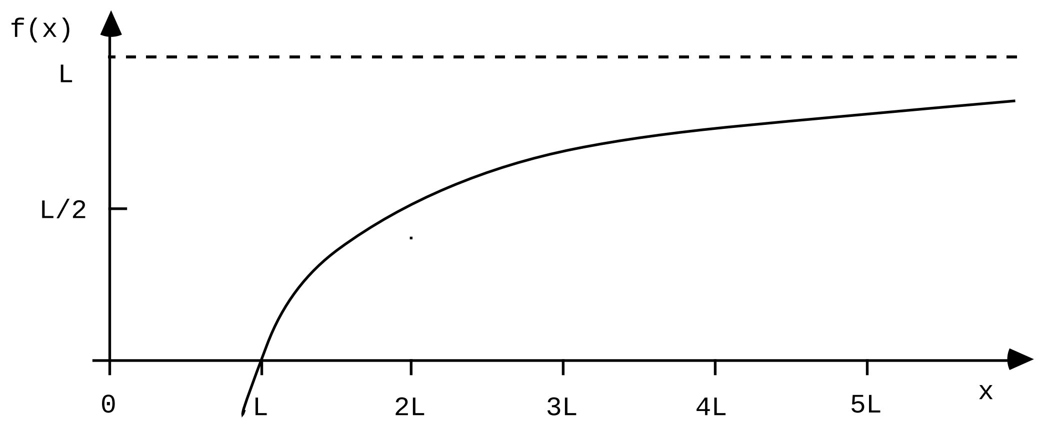

Fig. 2.3.1: Plot of the Welch Function $f(x) = \frac{L(x-L)}{x-1}$.

The most remarkable properties of this function $f(x)$ are, on one hand, that it reaches its horizontal asymptote quite fast with increasing argument x and, on the other hand, that the tangent has a steep slope $L/(L-1)$ at the zero-crossing point $x = L$. In the range $L \leq x \leq 2L$, the function is almost linear.

In the ESTEC CDMA application, we are interested in the relation between the size M of a code $\mathcal{C}$ and the smallness of the interference parameter $\tilde{\sigma}^2_{wc}(\Delta)$. According to ESTEC requirements (cf. [1]), the system must operate with small interuser interferences, i.e., $\tilde{\sigma}^2_{wc}(\Delta) \ll L$. For S-CDMA, this smallness condition for $\tilde{\sigma}^2_{wc}(0) = \sigma^2_{wc}(0)/(M-1)$ has the following implication on the size of a possible code as follows from the fact that the Welch bound (16) then gives $\tilde{\sigma}^2_{wc}(0) = f(M)$.

Conclusion on the Code Size for the Current ESTEC Application:

> An S-CDMA system, which requires small interuser interference, i.e., $\tilde{\sigma}^2_{wc}(0) \ll L$, can accommodate only about $M \approx L$ users.

In order to derive conclusions on $\tilde{\sigma}^2_{wc}(\Delta)$ from the Welch bound in the QS or in the asynchronous case, we require an assumption on the parameter $\tilde{\alpha}^2$, which measures per-user autocorrelation properties of the code sequence set. Assumptions on the smallness of $\tilde{\alpha}^2$ are justified because there is a high-priority demand that the **auto**correlation functions be good. Primary reasons for this demand come from requirements for synchronization and tracking of the code sequences. The exact form of the assumption on the smallness of $\tilde{\alpha}^2$ is not important as long as the number M of code sequences is large because, in (21), $\tilde{\alpha}^2$ is weighted with the small factor $\frac{\Delta}{(\Delta+1)M-1} \approx \frac{1}{M}$.

In case of A-CDMA, the relation between the size M of the code and the smallness of the interference parameter $\tilde{\sigma}^2_{wc}(L-1)$ is most simply illustrated by an example.

Example of A-CDMA with Few Users: Suppose, an A-CDMA system with sequences of length L accommodates $M = L/10$ users. Assume that the autocorrelation parameter takes on the conservatively large value of $\tilde{\alpha}^2 = L^2/100$. The new bound (23) still implies

$$\tilde{\sigma}^2_{wc}(L-1) > \frac{9}{10}L - 10$$

showing that $\tilde{\sigma}^2_{wc}(L-1)$ cannot be much smaller than L, as is required in the ESTEC system, for any number of users $M \geq 10$ as this gives $L = 10M \geq 100$.

Conclusion on the Choice of CDMA Systems for ESTEC:

> An A-CDMA system cannot fulfill the ESTEC requirements, i.e., there is no A-CDMA system with a large number of users that achieves the desired small value of the per-user crosscorrelation parameter $\tilde{\sigma}^2_{wc}(L-1)$, viz. $\tilde{\sigma}^2_{wc}(L-1) \ll L$. Thus, the CDMA system for the ESTEC application must be either an S-CDMA system or a QS-CDMA system with a small allowed phase offset Δ.

Among the well-known code sequence sets for CDMA, there are only the Hadamard code (see Appendix A) and the preferentially-phased Gold code that have $M \approx L$ code sequences and hence, that are candidates for the current ESTEC CDMA application. The preferentially-phased Gold code has very good autocorrelation properties, unlike the Hadamard code, and therefore, it is better suited for the ESTEC CDMA application. In the next Section, it will be shown that the preferentially-phased Gold code is essentially optimal for QS-CDMA, justifying the choice of this code.

We have investigated some alternative code sequence sets and found some new and nearly optimal for their size M. But they contain many more than L code sequences and thus they cannot satisfy the restrictive ESTEC requirement on the smallness of the interuser interference. A comparision of the code parameters of these alternative codes, such as the size M and the per-user interference $\tilde{\sigma}^2_{wc}(\Delta)$, to the corresponding parameters of the preferentially-phased Gold code is presented in Appendix A.

2.4 Essential Optimality of the Preferentially-Phased Gold Sequence Set

First we describe the preferentially-phased Gold sequence set. In order to have a convenient notation for the code sequences, we make use of the following one-to-one correspondence between binary sequences and ± 1-sequences. To every binary sequence $\mathbf{b} = [b_0, b_1, ..., b_{L-1}] \in GF(2)^L$, there is associated one and only one ± 1-sequence

$$\mathbf{x} = [x_0, x_1, ..., x_{L-1}],$$

where $x_i = (-1)^{b_i}$, $i = 0, 1, ..., L-1$. This correspondence will be denoted by $\mathbf{x} = e(\mathbf{b})$, the letter e reminding one of the exponentiation. By abuse of notation we shall often abreviate

$$C_{\mathbf{ab}}(k) \stackrel{def.}{=} C_{e(\mathbf{a})e(\mathbf{b})}(k)$$

and speak of the crosscorrelation between $\mathbf{a}$ and $\mathbf{b}$. The crosscorrelation value at the origin can be expressed in terms of Hamming weight $w_H(.)$ as

$$C_{\mathbf{ab}}(0) = L - 2w_H(\mathbf{a} + \mathbf{b}),$$

and, more generally, for any integer k,

$$C_{\mathbf{ab}}(k) = L - 2w_H(\mathbf{a} + T^k\mathbf{b}).$$

The Gold code set is most simply described in terms of a preferred pair $(\mathbf{u}, \mathbf{v})$ of binary maximal length (m-) sequences. The m-sequences have been extensively studied (see e.g. [5] or Vol. I of [7]). The name "maximal length" comes from the following basic relation to linear feedback shift registers (LFSR): An m-sequence is a sequence of the maximal period that can be achieved by a binary LFSR of some fixed length m starting in some non-zero state. The maximal period that a binary LFSR of length m can achieve is $L = 2^m - 1$; thus, all binary m-sequences have lengths $L = 2^m - 1$. One period

$$\mathbf{u} = [u_0, u_1, ..., u_{L-1}],$$

is what is usually meant as the m-sequence. The following basic properties of m-sequences will be of later use (cf. Chap. 5.4, Vol. I of [7]):

(i) The Hamming weight of an m-sequence $\mathbf{u}$ is $\frac{L+1}{2}$.

(ii) The **shift-and-add** property holds, i.e., for any integer i $(0 < i < L)$ there exists an integer l such that

$$\mathbf{u} + T^i\mathbf{u} = T^l\mathbf{u}.$$

Furthermore, if i runs though $1, 2, ..., L-1$ then so does l, but in some different order.

Let $L = 2^m - 1$ and set $e = \lfloor m+2 \rfloor / 2$. A **preferred pair** $(\mathbf{u}, \mathbf{v})$ **of sequences of length** $2^m - 1$ is obtained by starting with an m-sequence $\mathbf{u}$ and taking $\mathbf{v}$ to be the $(2^e + 1)$st decimation of $\mathbf{u}$, i.e., the i-th component of $\mathbf{v}$ is given by $v_i = u_{i(2^e+1)}$. The sequence $\mathbf{v}$ is also an m-sequence unless and only unless m is divisible by 4. When m is not divisible by 4, we will call $(\mathbf{u}, \mathbf{v})$ a **preferred pair of m-sequences**. A preferred pair of sequences of length $2^m - 1$ has a correlation function $C_{\mathbf{uv}}(k)$ that assumes only the three values $-1, -1 + 2^{(m+g)/2}$ and $-1 - 2^{(m+g)/2}$, where $g = gcd(2e, m)$ is the greatest common divisor of $2e$ and m. The multiplicities of the three values are given in the following table (see, e.g., Corollary 11.14 in [8]) for the case when $(\mathbf{u}, \mathbf{v})$ is a preferred pair of m-sequences:

$C_{\mathbf{ab}}(k)$	Multiplicity
-1	$2^m - 2^{m-g} - 1$
$-1 + 2^{(m+g)/2}$	$2^{m-g-1} + 2^{(m-g)/2-1}$
$-1 - 2^{(m+g)/2}$	$2^{m-g-1} - 2^{(m-g)/2-1}$

(22)

Table 2.4.1: Histogram of the Correlation Function of a Preferred Pair of m-Sequences.

Now we are ready to define the **preferentially-phased Gold code set** $\mathcal{G}$. Let $(\mathbf{u}, \mathbf{v})$ be a preferred pair of sequences of length $2^m - 1$, then

$$\mathcal{G} \stackrel{def.}{=} \{\mathbf{v}\} \cup \{\mathbf{v} + T^i\mathbf{u} | i = 0, 1, ..., L-1\}.$$

[Hereafter, we will restrict our analysis to the case where m is not divisible by 4 or, equivalently, where $\mathbf{v}$ is an m-sequence. Results for the case m divisible by 4 will be stated without proof. If m is divisible by 4, the non-m-sequence $\mathbf{v}$ should be removed from the preferentially-phased Gold set, if one wishes to retain the good **autocorrelation** properties of all sequences in the set.] The cardinality of the preferentially-phased Gold code set $\mathcal{G}$ is $M = L + 1$. The set of all ± 1-sequences $e(\mathbf{a}), \mathbf{a} \in \mathcal{G}$, will by abuse of notation again be denoted by $\mathcal{G}$. [The full Gold code sequence set of length L contains $L+2$ sequences. In the preferentially-phased Gold sequence set, we have omitted the sequence $\mathbf{u}$ that would be in the corresponding full Gold sequence set because the optimal crosscorrelation Property 1 below would not hold for this full Gold sequence set. Obtaining this optimal correlation at the origin is, for S-CDMA, well worth the reduction of the number of possible users M from $L+2$ to $L+1$.] For any two sequences $\mathbf{x} \neq \mathbf{y}$ in $\mathcal{G}$, say, $\mathbf{x} = e(\mathbf{v}+T^i\mathbf{u})$, $\mathbf{y} = e(\mathbf{v}+T^j\mathbf{u})$, one has

$$\begin{aligned} C_{\mathbf{xy}}(0) &= L - 2w_H(T^i\mathbf{u} + T^j\mathbf{u}) = L - 2w_H(\mathbf{u} + T^{j-i}\mathbf{u}) \\ &= L - 2w_H(T^l\mathbf{u}) = L - 2w_H(\mathbf{u}) = -1. \end{aligned}$$

In the case $\mathbf{x} = e(\mathbf{v})$, the same crosscorrelation value occurs. Note that $|C_{\mathbf{xy}}(0)| = 1$ is the smallest possible crosscorrelation value because the length L is odd. Thus, we have the following optimal crosscorrelation property of the Gold code $\mathcal{G}$ at the origin.

Property 1: $C_{\mathbf{xy}}(0) = -1$, for all $\mathbf{x}, \mathbf{y} \in \mathcal{G}, \mathbf{x} \neq \mathbf{y}$.

Property 1 implies that the Welch bound (16) is satisfied with equality. The crucial parameter σ_{wc}^2 of Criterion 1 takes on its minimum possible value on a per-user basis and therefore it follows that the preferentially-phased Gold code set $\mathcal{G}$ is optimal for S-CDMA.

Conclusion on the Choice of Code for S-CDMA:

> The preferentially-phased Gold code set $\mathcal{G}$ is optimal for S-CDMA with respect to Criterion 1 for $M = L + 1$ users and gives $\tilde{\sigma}^2_{wc} = 1$.

Using the preferentially-phased Gold code set $\mathcal{G}$ in QS-CDMA with a maximal phase offset Δ, $0 < \Delta \leq 1$, we are interested in the worst-user worst-case interuser interference $\max_i \max_{\boldsymbol{\delta}} \frac{1}{2}\sigma_i^2(\boldsymbol{\delta})$ or in the worst-user average interuser interference $\sigma^2 = \max_i \frac{1}{2} E_{\boldsymbol{\delta}}[\sigma_i^2(\boldsymbol{\delta})]$, depending on whether the phase offsets are slowly or rapidly varying, respectively. Because the preferentially-phased Gold code $\mathcal{G}$ has its smallest crosscorrelation values at the origin, it follows that $|C_{\mathbf{xy}}(1)| \geq |C_{\mathbf{xy}}(0)| = 1$ for any two different preferentially-phased Gold sequences $\mathbf{x}, \mathbf{y}$, and therefore, that the worst-user worst-case interference for $\Delta = 1$ is given by

$$\sigma^2_{wc}(1) = \max_{\mathbf{x} \in \mathcal{G}} \sum_{\substack{\mathbf{y} \in \mathcal{G} \\ \mathbf{y} \neq \mathbf{x}}} C^2_{\mathbf{xy}}(1).$$

Note that by Criterion 2 this is the crucial parameter one wishes to minimize. We shall now show that the value for $\sigma^2_{wc}(1)$, as obtained by the preferentially-phased Gold code set $\mathcal{G}$, cannot be substantially reduced by choosing another code $\mathcal{C}$ with $M = L + 1$ code sequences if one requires that this other code $\mathcal{C}$ is also optimal for S-CDMA and that $\mathcal{C}$ has the same or smaller worst-case off-peak autocorrelation values $< \mathbf{x}, T\mathbf{x} >^2$, $\mathbf{x} \in \mathcal{C}$, as the preferentially-phased Gold code $\mathcal{G}$. Such an alternative code $\mathcal{C}$ will be called a **competitor code**. Thus, we shall show that the preferentially-phased Gold code $\mathcal{G}$ is almost optimal for QS-CDMA among all its competitor codes; this will be referred to as the **essential optimality** of the preferentially-phased Gold code set $\mathcal{G}$ for QS-CDMA.

The argument for the essential optimality of the preferentially-phased Gold code $\mathcal{G}$ for QS-CDMA relies on the Welch bound; in particular, it is the condition for equality of the Welch bound and its interpretation given by the array (15) that will be exploited. This permits us to bound the sum of $C^2_{\mathbf{xy}}(1)$ over all code sequences $\mathbf{x}$ and $\mathbf{y}$ without using any knowledge of the individual terms in the sum. Because the preferentially-phased Gold code $\mathcal{G}$ satisfies the Welch bound (16) with equality, it follows from the Condition for Equality in the Welch Bound that the columns of the array (16) are orthogonal. We now consider the associated virtual code set $\mathcal{G}_\Delta$ with $\Delta = 1$ and we write out its code sequences as rows in the following array

$$\begin{array}{cccccc}
x_0^{(1)} & x_1^{(1)} & \cdot\ \cdot\ \cdot & & \cdot & x_{L-1}^{(1)} \\
x_0^{(2)} & x_1^{(2)} & \cdot\ \cdot\ \cdot & & \cdot & x_{L-1}^{(2)} \\
\vdots & \vdots & & & & \vdots \\
x_0^{(M)} & x_1^{(M)} & \cdot\ \cdot\ \cdot & & \cdot & x_{L-1}^{(M)} \\
x_1^{(1)} & x_2^{(1)} & \cdot\ \cdot\ \cdot & & x_{L-1}^{(1)} & x_0^{(1)} \\
x_1^{(2)} & x_2^{(2)} & \cdot\ \cdot\ \cdot & & x_{L-1}^{(2)} & x_0^{(2)} \\
\vdots & \vdots & & & \vdots & \vdots \\
x_1^{(M)} & x_2^{(M)} & \cdot\ \cdot\ \cdot & & x_{L-1}^{(M)} & x_0^{(M)}.
\end{array} \tag{23}$$

Note that the L subcolumns formed by the last M rows are a cyclic permutation of the L subcolumns given by the first M rows. Because the upper L subcolumns are orthogonal, so are the L lower subcolumns; hence, all columns of the array are orthogonal. This is precisely the condition for equality in the Welch bound when applied to the virtual preferentially-phased Gold code set $\mathcal{G}_\Delta$. Thus, the Welch bound holds with equality and one has

$$\sum_{\mathbf{x}\in\mathcal{G}_\Delta}\sum_{\mathbf{y}\in\mathcal{G}_\Delta} < \mathbf{x},\mathbf{y} >^2 = (2M)^2 L. \tag{24}$$

Using symmetry properties of the double sum and of the scalar product, the left side can be written as

$$\begin{aligned} & \sum_{n=1}^{M}\left(\sum_{m=1}^{M} < \mathbf{x}^{(n)},\mathbf{x}^{(m)} >^2 + \sum_{m=1}^{M} < \mathbf{x}^{(n)},T\mathbf{x}^{(m)} >^2\right) + \\ & \sum_{n=1}^{M}\left(\sum_{m=1}^{M} < T\mathbf{x}^{(n)},\mathbf{x}^{(m)} >^2 + \sum_{m=1}^{M} < T\mathbf{x}^{(n)},T\mathbf{x}^{(m)} >^2\right) \\ = \; & 2\sum_{n=1}^{M}\left(\sum_{m=1}^{M} < \mathbf{x}^{(n)},\mathbf{x}^{(m)} >^2 + \sum_{m=1}^{M} < \mathbf{x}^{(n)},T\mathbf{x}^{(m)} >^2\right). \end{aligned}$$

The preferentially-phased Gold code set $\mathcal{G}$ satisfies the Welch bound with equality, therefore $\sum_{n=1}^{M}\sum_{m=1}^{M} < \mathbf{x}^{(n)},\mathbf{x}^{(m)} >^2 = M^2 L$ and (24) becomes

$$2M^2L + 2\sum_{n=1}^{M}\sum_{m=1}^{M} < \mathbf{x}^{(n)},T\mathbf{x}^{(m)} >^2 = 4M^2L$$

or equivalently

$$\sum_{n=1}^{M}\sum_{m=1}^{M} < \mathbf{x}^{(n)},T\mathbf{x}^{(m)} >^2 = M^2 L. \tag{25}$$

Note that this equation also holds for every code sequence set $\mathcal{C}$, which satisfies the Welch bound with equality, regardless of the autocorrelation properties of the code. In particular, it holds for a competitor code.

We shall now derive a lower bound on $\sigma_{wc}^2(1)$ that will hold for every competitor code $\mathcal{C}$. Moving the autocorrelation terms of the equation above to the right side, we get

$$\sum_{n=1}^{M}\sum_{m\neq n} < \mathbf{x}^{(n)},T\mathbf{x}^{(m)} >^2 = M^2L - \sum_{n=1}^{M} < \mathbf{x}^{(n)},T\mathbf{x}^{(n)} >^2 . \tag{26}$$

The following Lemma will help to evaluate the right side of this equation.

Lemma 2: Let $(\mathbf{u},\mathbf{v})$ be a preferred pair of m-sequences and let $\mathcal{G}$ denote the associated preferentially-phased Gold code set. Then, one has

(i) $\sum_{\mathbf{x}\in\mathcal{G}} < \mathbf{x},T\mathbf{x} >^2 = 1 + \sum_{k=0}^{L-1} C_{\mathbf{uv}}^2(k)$.

(ii) $\sum_{k=0}^{L-1} C_{\mathbf{uv}}^2(k) = L^2 + L - 1$.

(iii) $\sum_{k=0}^{L-1} C_{\mathbf{uv}}(k) = -1$.

Proof. Every code sequence $\mathbf{x} \neq e(\mathbf{v})$ can be written as $\mathbf{x} = e(\mathbf{v} + T^i\mathbf{u}), i = 0, 1, ..., L-1$, yielding $< \mathbf{x}, T\mathbf{x} >= C_{\mathbf{x},T\mathbf{x}}(0) = L - 2w_H(\mathbf{v} + T\mathbf{v} + T^i\mathbf{u} + T^{i+1}\mathbf{u})$. By the Shift-and-Add Property of m-sequences, one has $\mathbf{v} + T\mathbf{v} = T^s\mathbf{v}$ for some s and $\mathbf{u} + T\mathbf{u} = T^t\mathbf{u}$ for some t. If i runs through $0, 1, ..., L-1$, so does $j \equiv t + i (\mathrm{mod} L)$. Thus, one obtains

$$\sum_{i=0}^{L-1} < e(\mathbf{v} + T^i\mathbf{u}), Te(\mathbf{v} + T^i\mathbf{u}) >^2 = \sum_{j=0}^{L-1} C^2_{T^s\mathbf{v}T^j\mathbf{u}}(0) = \sum_{j=0}^{L-1} C^2_{\mathbf{vu}}(j-s) = \sum_{k=0}^{L-1} C^2_{\mathbf{uv}}(k).$$

The only remaining code sequence $\mathbf{x} = e(\mathbf{v})$ has the autocorrelation value

$$< e(\mathbf{v}), Te(\mathbf{v}) >= L - 2w_H(T^s\mathbf{v}) = L - 2w_H(\mathbf{v}) = -1.$$

This completes the proof for (i).

To prove (ii) and (iii), one simply sums up all terms using Table 2.4.1.

Applying (i) and (ii) to the right side of (26), this equality becomes

$$\sum_{n=1}^{M} \sum_{m \neq n} < \mathbf{x}^{(n)}, T\mathbf{x}^{(m)} >^2 = M^2L - ML = M(M-1)L \tag{27}$$

and the worst-user worst-case interuser interference $\sigma^2_{wc}(1)$ can be bounded by the average, i.e.,

$$\sigma^2_{wc}(1) \geq \frac{M(M-1)L}{M} = L^2. \tag{28}$$

Property 2: The worst-user worst-case interference $\sigma^2_{wc}(1)$ of any competitor code $\mathcal{C}$ is lower bounded by L^2.

Remark: It can be shown that (27) and (28) also hold for the case when m is divisible by 4, i.e., when $\mathbf{v}$ is not an m-sequence. Hence, the preferentially-phased Gold code set $\mathcal{G}$ satisfies Properties 2 and 3 also for the case, when m is divisible by 4.

We shall give an interpretation of equation (27) in terms of the average-user worst-case interference

$$\sigma^2_{avg} \stackrel{def.}{=} \frac{1}{M} \sum_{n=1}^{M} \sum_{m \neq n} < \mathbf{x}^{(n)}, T\mathbf{x}^{(m)} >^2 .$$

Property 3: The average-user worst-case interference σ^2_{avg} of a competitor code $\mathcal{C}$ for QS-CDMA ($\Delta = 1$) is lower bounded by the average-user worst-case interference L^2 of the preferentially-phased Gold code $\mathcal{G}$, viz.

$$\sigma^2_{avg} \geq L^2.$$

We shall show hereafter that the worst-user worst-case interference for the preferentially-phased Gold code set $\mathcal{G}$ equals $\sigma^2_{wc}(1) = L^2 + L - 1$. Thus, the actual value of the worst-user worst-case interference for the preferentially-phased Gold code set $\mathcal{G}$ is only slightly larger than the general lower bound (28), which is valid for any competitor code $\mathcal{C}$. A substantial reduction on $\sigma^2_{wc}(1)$ could only be achieved if the alternative code $\mathcal{C}$ had very large autocorrelation values $< \mathbf{x}, T\mathbf{x} > \approx L^2$ for all $\mathbf{x} \in \mathcal{C}$. But such large autocorrelation values are in general not tolerable and, in particular, a competitor code has much smaller autocorrelation values. From Properties 2 and 3, one can draw the following conclusion on the suitability of the preferentially-phased Gold code $\mathcal{G}$ for QS-CDMA.

Conclusion on the Choice of Code for QS-CDMA:

> The preferentially-phased Gold code $\mathcal{G}$ is essentially optimal with respect to Criterion 2 among all competitor codes.
>
> Taking the smallness of the average-user worst-case interference σ^2_{avg} as the optimizing criterion for code selection, it follows that the preferentially-phased Gold code $\mathcal{G}$ is optimal among all competitor codes.

Remarks: 1. The code $\mathcal{C}$ obtained from the preferentially-phased Gold code $\mathcal{G}$ by choosing different phases of the code sequences is a competitor code, provided it satisfies Property 1. By Property 2 it follows that such a code $\mathcal{C}$ could have only a slightly smaller $\sigma^2_{wc}(1)$ than the preferentially-phased Gold code $\mathcal{G}$. We conjecture that, if it ever happens that any other choice of phases produces an optimal code for S-CDMA, then the worst-user worst-case interference $\sigma^2_{wc}(1)$ will be the same as for the preferentially-phased Gold code set $\mathcal{G}$, i.e., that no reduction of the interference $\sigma^2_{wc}(1)$ is possible by choosing different phases of the preferentially-phased Gold sequences. Unfortunately, we could not prove this statement, which is more of theoretical than of practical interest.
2. Should the ESTEC requirements on the smallness of the crosscorrelation functions at the origin change, the reader is refered to Appendix A for a thorough discussion of alternative virtually optimal code sequence sets. For code sequence sets satisfying the Welch bound with equality, (25) and (26) hold. For such code sequence sets, one can derive a lower bound like (28) on the worst-user worst-case interuser interference $\sigma^2_{wc}(1)$ as done above. This lower bound allows one to bound the interuser interference $\sigma^2_{wc}(1)$ of any competitor code competing with the particular chosen code sequence set.

It remains to determine the value of the worst-user worst-case interference $\sigma^2_{wc}(1)$ of the preferentially-phased Gold code $\mathcal{G}$. We start with some useful identities, which can be derived by an argument similar to the proof of Lemma 2; they are collected in the next Lemma.

Lemma 3: Let $(\mathbf{u}, \mathbf{v})$ be a preferred pair of m-sequences and let $\mathcal{G}$ denote the associated preferentially-phased Gold code set. Then, one has

(i) $\sum_{\mathbf{y} \in \mathcal{G}} < \mathbf{x}, T\mathbf{y} >^2 = 1 + \sum_{k=0}^{L-1} C^2_{\mathbf{uv}}(k) = L^2 + L$, for all $\mathbf{x} \in \mathcal{G}$;

(ii) $\sum_{\mathbf{y} \in \mathcal{G}} < \mathbf{x}, T\mathbf{y} > = -1 + \sum_{k=0}^{L-1} C_{\mathbf{uv}}(k) = -2$, for all $\mathbf{x} \in \mathcal{G}$.

Applying Lemma 3 (i) to the Welch bound as given by (25), one sees that the inner sum does not depend on the particular code sequence $\mathbf{x}^{(n)}$, thus, one gets

$$\sum_{m=1}^{M} < \mathbf{x}^{(n)}, T\mathbf{x}^{(m)} >^2 = \frac{LM^2}{M} = LM = L(L+1) \text{ for all } n = 1, 2, ..., M,$$

or equivalently

$$\sum_{m\neq n} < \mathbf{x}^{(n)}, T\mathbf{x}^{(m)} >^2 = L^2 + L - < \mathbf{x}^{(n)}, T\mathbf{x}^{(n)} >^2 \quad \text{for all } n = 1, 2, ..., M.$$

This gives a complete description of the interuser interference for all users. [One can show, using a slightly different argument, that this same description of the interuser interference holds also for the case when m is divisible by 4.] The worst case is obtained when $< \mathbf{x}^{(n)}, T\mathbf{x}^{(n)} >^2 = 1$, thus, one has

$$\sigma_{wc}^2(1) = L^2 + L - 1.$$

We consider a QS-CDMA system using the preferentially-phased Gold code $\mathcal{G}$ and having a maximal allowed phase offset Δ, $0 < \Delta \leq 1$ as proposed in the ESTEC CDMA application. We are now able to determine the crucial parameters $\max_i \max_{\boldsymbol{\delta}} \frac{1}{2}\sigma_i^2(\boldsymbol{\delta})$ and $\frac{1}{2}E_{\boldsymbol{\delta}}[\sigma_i^2(\boldsymbol{\delta})]$ for slowly-varying and rapidly-varying phase offsets.

In the case of slowly-varying phase offsets the worst-user worst-case interference is determined by $\sigma_{wc}^2(1)$ according to the derivation of Criterion 2, i.e., one obtains

$$\max_i \max_{\boldsymbol{\delta}} \frac{1}{2}\sigma_i^2(\boldsymbol{\delta}) = \frac{1}{2}\Delta^2\sigma_{wc}^2(1) = \frac{1}{2}\Delta^2(L^2 + L - 1) \tag{29}$$

whenever $\frac{1}{\sqrt{L}} < \Delta \leq 1$. Note that, for $\Delta \leq \frac{1}{\sqrt{L}}$, the crosscorrelation values at the origin will determine the interuser interference and one then has $\max_i \max_{\boldsymbol{\delta}} \frac{1}{2}\sigma_i^2(\boldsymbol{\delta}) = \frac{1}{2}L$; hence this very improbable case coincides with S-CDMA.

In the case of rapidly-varying phase offsets, the worst-user case occurs for a user with optimal autocorrelation value $< \mathbf{x}^{(i)}, T\mathbf{x}^{(i)} >^2 = 1$ because in this case the positive definite quadratic form given by (11) is maximal. Using Property 1, Lemma 2 and Lemma 3, one can derive

$$\begin{aligned} \max_i \frac{1}{2}E[\sigma_i^2(\boldsymbol{\delta})] &= \frac{1}{2}\frac{\Delta^2}{12}(2(L^2+L-1)) + \frac{1}{2}(1 - \frac{2\Delta}{3} + \frac{\Delta^2}{6})L + \frac{1}{2}(\frac{\Delta}{3} - \frac{\Delta^2}{6})(-1)(-1-1) \\ &= \frac{\Delta^2}{12}(L^2+L-1) + (\frac{1}{2} - \frac{\Delta}{3} + \frac{\Delta^2}{12})L + \frac{\Delta}{3} - \frac{\Delta^2}{6}. \end{aligned} \tag{30}$$

[For the case when m is divisible by 4, the variance of the interuser interference could not be determined in a closed form like (30), but one can show that $\max_i \frac{1}{2}E[\sigma_i^2(\boldsymbol{\delta})] \approx \frac{\Delta^2}{12}L^2$, neglecting terms that are of order smaller than 2.] Comparing slowly-varying and rapidly-varying phase offsets, one notices that the interference power of the rapidly-varying case is about 6 times (7.8 dB) smaller than for the slowly-varying case. Thus, it is advantageous to have a system with rapidly-varying phase offsets; furthermore, as the formula for the worst-user average interference has a quadratic dependence on Δ, it is desirable to have Δ as small as possible.

Crucial Performance Parameters for the ESTEC QS-CDMA System:

For the ESTEC QS-CDMA system relying on the preferentially-phased Gold code $\mathcal{G}$, the value of the interuser interference parameter is given by (29) for slowly-varying phase offsets, and by (30) for rapidly-varying phase offsets.

3 Relation between Spectrum Spreading and Coding

3.1 The Coding/Spreading Problem

Because a binary error-correcting code of rate R (measured in information bits per encoded bit) already expands the "bandwith" of the information bit stream by a factor of $1/R$, the question naturally arises as to how much of the ultimate spreading of the spectrum should be done by coding and how much by the use of spreading sequences, or whether in fact these two spreadings should somehow be done more or less independently.

The conceptually simplest coding/spreading option is the **Cascade of Coding with Ordinary CDMA**: The "data symbols" $\{b_j\}$ considered in Section 2 of this report are then the encoded bits from the rate R encoder; each encoded bit b_j determines the polarity of the length L spreading sequence $\mathbf{x}$.

In this option, the overall spreading factor is L/R. The case $R = 1$ corresponds to uncoded signalling, while the case $L = 1$ corresponds to unspread transmission of the encoded sequence. The obvious question is how to choose the parameters L and R to obtain the best system.

The less obvious option, favored by Viterbi [12], will here be called the **Superposition of User-Specific Spreading on the Coded Sequence**: The "spreading sequence" for each user is now a periodic sequence with very long period T; the "data symbols" $\{b_j\}$ are again the encoded bits from the rate R encoder, but each encoded bit b_j now determines the polarity of only the j-th **subsequence** of length L (where $L \ll T$) within the spreading sequence. Again the overall spreading factor is L/R. In fact, this is none other than conventional "direct-sequence spreading" of the encoded signal by the periodic sequence — the new twist is that the direct-sequence spreading is being used to gain a multiple-access capability.

When superposition of user specific spreading on the coded sequences of each user is employed to gain a multiple-access capability, a certain amount of coordination (or synchronization) of the users is required. First, the receiver must know the delay (relative to the system clock) of the periodic sequence that is used for direct-sequence spreading by user i of his coded sequence in order that the receiver can appropriately despread. Second, these delays must be coordinated so that for every pair of users they are always seperated by at least one chip of the periodic sequence in order to prevent large steady interuser interference.

To gain insight into these two options, we first consider their properties for a single user before considering the case of actual interest where the number M of users is large. For conceptual simplicity, we will consider a baseband spread-spectrum system, but the results will be seen to generalize obviously to a bandpass system. Because a convolutional encoding system with a soft-decision Viterbi decoder is virtually certain to be the choice of coding system in an actual implementation by ESTEC, we restrict ourselves to such systems.

3.2 Coding/Spreading for a Single User

For the single-user baseband system, we may assume binary antipodal modulation is used. In this case, the **asymptotic coding gain** (ACG), which is the usual parameter used to compare coded systems that must operate at small post-decoding error rates (say 10^{-5} or less), of the convolutional coding system on an unspread AWGN channel is given by

$$ACG = Rd_{free} \tag{31}$$

where d_{free} is the free distance of the convolutional code, i.e., the minimum Hamming distance between encoded paths that diverge from any state in the code trellis. The explanation of (31) is as follows.

Compared to no coding at the same signal power, coding has reduced the energy of each transmitted bit by a factor of R but has increased the squared Euclidean distance between competing signals that can cause an error by a factor of d_{free}.

The ACG will be a good measure of the power advantage purchased by coding provided that

(i) the number of competing paths in the code trellis at distance d_{free} from any given path is not too large and

(ii) the signal-to-noise ratio γ is large enough that decoding errors occur with high probability only to these closest competing paths, which is roughly equivalent to demanding a post-decoding error rate of about 10^{-5} or less.

A surprising fact (in light of much existing literature although it was stated rather clearly in [13]) is the **Principle of Asymptotic-Coding-Gain Invariance:** *On a single-user AWGN channel, the asymptotic coding gain (ACG) is not changed by spreading*, regardless of whether this spreading is done by cascade of the encoder with a spreading sequence of length L or whether this spreading is done by superposition of spreading on the coded sequence in the manner that each encoded bit controls the polarity of an L-bit segment of a spreading sequence with period $T \gg L$.

The reason is that in both cases the effect of spreading can be viewed as transforming the code trellis to a new code trellis in which each encoded bit b_j on a trellis branch at some fixed depth into the trellis is replaced by an L bit sequence $b\mathbf{x}$. Because the Hamming distance between $b\mathbf{x}$ and $-b\mathbf{x}$ is L, d_{free} is increased by a factor of L for this new trellis code, but the rate of this new code is decreased by a factor of $1/L$ since there are now L times as many "encoded bits" per information bit. Thus the product $Rd_{free} = ACG$ is unchanged.

Other effects, such as multipath propagation that negates the AWGN model, may make it desirable to spread the coded signal, but one should not deceive himself into thinking that spreading either increases or decreases the coding gain on a single-user AWGN channel. (The so-called "processing gain" of a spread-spectrum system is an entirely misleading name when used for a single-user AWGN-channel system.)

We have not yet introduced the other important parameter of the coding system, its **decoding complexity** μ, which we define as the base-two logarithm of the number of encoded states, i.e., there are 2^{μ} encoded states. Because the complexity of a Viterbi decoder is proportional to this number of states, it is vital that μ be small: $\mu = 7$ is near the practical maximum.

3.3 Uniform, Orthogonal, Superorthogonal and EPUM Codes

Uniform codes were introduced by Massey in 1963. The binary uniform codes can be described (cf. [14]) in terms of Viterbi's parameter K, which is the number of information bits in the constraint span of the code, as the codes that can be encoded in the manner shown in Fig. 3.3.1

where the "periodic $(K-1)$-tuple generator" is any device (say, a $(K-1)$-bit ring counter) that generates periodically in any specified order the 2^{K-1} different binary $(K-1)$-tuples and where the information-bit register is shifted at the beginning of each cycle only. Thus, the rate is $R = 1/2^{K-1}$, and it can easily be shown (cf. [14]) that $d_{free} = (K+1)2^{K-2}$. The complexity of the code is $\mu = K-1$ because the leftmost information bit in the information-bit register of Fig. 3.3.1 is the "current input" to the encoder and only the remaining $K-1$ bits are the "current state" of the encoder (i.e., they constitute what must be "remembered" about past inputs). Thus, the asymptotic coding gain of the uniform codes is

$$ACG_{unif} = \frac{1}{2}(K+1) = \mu/2 + 1 \tag{32}$$

for any μ, $\mu = 1, 2, 3, \ldots$.

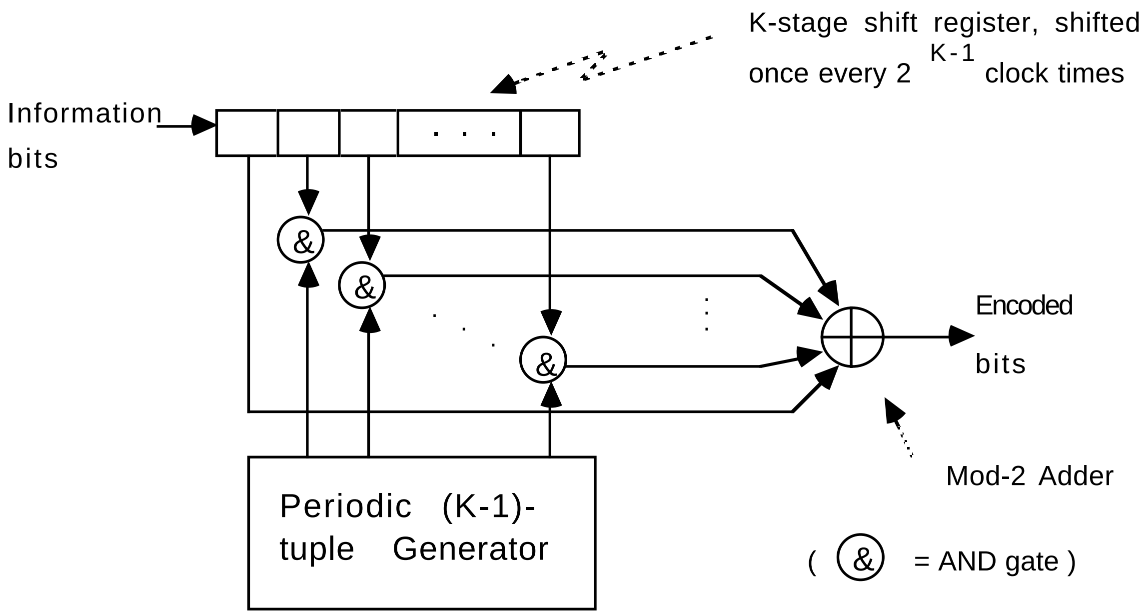

Fig. 3.3.1: Encoder for a Uniform Code

Orthogonal convolutional codes were introduced by Viterbi in 1967 (cf. [15, pp. 252–258]). The encoder for the uniform code in Fig. 3.3.1 becomes an encoder for an orthogonal code if the leftmost stage of the information-bit register (and its associated output line to the mod-2 Adder) is removed. (The K and μ of the orthogonal code are thus both one less than for the corresponding uniform code.) The code parameters are $R = 1/2^K$, $d_{free} = K2^{K-1}$ and $\mu = K-1$ so that the asymptotic coding gain is

$$ACG_{orthog} = \frac{1}{2}K = \mu/2 + 1/2. \tag{33}$$

One sees from (32) and (33) that, for a given decoding complexity μ, the orthogonal code is slightly inferior to the uniform code.

Superorthogonal codes were recently introduced by Viterbi for use in QUALCOMM's innovative CDMA system. The encoder of Fig. 3.3.1 becomes an encoder for a superorthogonal code if the information-bit register is augmented with an additional stage on the right whose output is also fed to the mod-2 adder. (The K and μ of the superorthogonal code are thus both one greater than for the corresponding uniform code.) The code parameters are $R = 1/2^{K-2}$, $d_{free} = (K+2)2^{K-3}$ and $\mu = K - 1$. Thus, the asymptotic coding gain is

$$ACG_{super} = \frac{1}{2}(K+2) = \mu/2 + 3/2. \tag{34}$$

One sees from (32) and (34) that, for a given decoding complexity μ, the superorthogonal code is slightly superior to the uniform code.

Equidistant Partial-Unit Memory (EPUM) codes were introduced by Lauer in 1979 [16]. For any μ ($\mu = 1, 2, 3, ...$) there is an EPUM of rate $R = 1/2^\mu$ with $d_{free} = (\mu+1)2^\mu$ [there are also EPUM's with other less convenient parameters] and hence with asymptotic coding gain

$$ACG_{EPUM} = \mu + 1 \tag{35}$$

[which expression holds for all EPUM's]. For large μ, we see that the ACG of EPUM's is $3dB$ better than for uniform, orthogonal or superorthogonal codes. For realistic μ, the gain is smaller compared to the superorthogonal codes; for instance with $\mu = 5$ (32 encoder states), the ACG of 6 for the EPUM is only $1.76dB$ better than the ACG of 4 for the superorthogonal code.

The "catch" of the EPUM's is that they are not "conventional" convolutional codes of rate $R = 1/n$, where by the latter we mean that one information bit enters and n encoded bits leave the encoder in each basic encoding step. For the EPUM of rate $R = 1/2^\mu$, $\mu+1$ information bits enter and $(\mu+1)2^\mu$ encoded bits leave the encoder in each basic encoding step. This means that each state in the code trellis has 2^μ times as many successors as for a conventional code of the same **rate** (regardless of the latter's decoding complexity). This certainly adds to the "true complexity" of the corresponding Viterbi decoder compared to a conventional code of the same **decoding complexity** μ (regardless of the latter's rate), but the fact that the Viterbi decoder for the EPUM decodes $\mu+1$ information bits at each decoding step somewhat mitigates this disadvantage. The "right way" to compare an EPUM to, say, a superorthogonal code is on the basis of **asymptotic coding gain** as will be illustrated by an example. The EPUM with $\mu = 2$ (which has $R = 1/4$) gives the same asymptotic coding gain ($3dB$) as the superorthogonal code with $\mu = 5$ (which has $R = 1/16$). By any reasonable measure, the Viterbi decoder for the EPUM trellis that has only 4 states is simpler than the Viterbi decoder for the superorthogonal trellis that has 32 states, even though each state in the former trellis has 8 successors (there are two parallel transitions to each of the four distinct successor states) whereas each state in the latter trellis has only 2 successors. When used in a spread-spectrum system, the same desired large overall spreading factor of L/R can be achieved with both codes by appropriate choice of L, as was pointed out in section 3.1 above.

The purpose of the above discussion was to show that the use of superorthogonal convolutional codes provides no new breakthrough in coding for a spread-spectrum system. Superorthogonal codes offer only slightly better asymptotic coding gain than the venerable uniform codes of the same true decoding complexity, and are arguably somewhat inferior to the EPUM codes that have been known for over a decade (but never exploited in practice to our knowledge). This does not mean that superorthogonal codes are not useful. The fact that the branch metrics for these codes can be quickly computed is an important implementation advantage of these codes. Our point here is that these codes provide ACG's only slightly better than codes that have long been known.

3.4 Coding/Spreading for Many Users

For the many-user (baseband) spread-spectrum system, the coding discussion of the preceding sections applies unchanged for each user except that (on the reasonable assumption that the Gaussian approximation for the inter-user interference is valid) the variance of the inter-user interference must now be taken into account as additional Gaussian noise. If the spreading is done by the **cascade of coding with ordinary CDMA** as described in section 3.1, then all of the analysis of section 2 of this report applies unchanged for the calculation of this variance regardless of whether S-CDMA or QS-CDMA is chosen. The advantages, if any, of S-CDMA and QS-CDMA over A-CDMA are also maintained without change.

The interesting case is when the spreading is done by **superposition of user-specific spreading on the coded sequences**, which was also described in section 3.1 and which (as will be seen) is the real breakthrough of the QUALCOMM system when used with very many users. We consider now such spreading by superposition.

Consider first the case of S-CDMA, i.e., the case when the length L subsequences (that are polarity-modulated by each user) are perfectly aligned at the receiver. On the reasonable assumptions that

(i) the spreading sequence is a periodic maximal-length sequence (or "m-sequence") of period $T = 2^L - 1$ and hence all L chip segments of this sequence are distinct and each possible L chip pattern (except $\mathbf{0}$) appears somewhere in the sequence and

(ii) the segment used at any time by each user is spaced by at least 1 chip position from that of every other user,

[which guarantees that the segments $\mathbf{x}$ and $\mathbf{y}$ used by any two users at any time are both different and enjoy virtual time-statistical independence (because of the "shift-and-add property" of m-sequences) when $\gcd(L, 2^L - 1) = 1$] then the inter-user interference term $< \mathbf{x}, \mathbf{y} >$ seen by the $\mathbf{x}$-segment user has a (time-average) mean of virtually 0 when L is large (say $L \geq 32$) and has a (time-average) second moment, and hence virtually variance, given by

$$E[< \mathbf{x}, \mathbf{y} >^2] = \frac{1}{2^L - 1} \sum_{i=1}^{L} (L - 2i)^2 \begin{pmatrix} L \\ i \end{pmatrix}.$$

Thus, this per-user inter-user interference has a (time-average) variance of

$$\begin{aligned} \tilde{\sigma}^2 &\approx \frac{1}{2^L} \sum_{i=0}^{L} (L - 2i)^2 \begin{pmatrix} L \\ i \end{pmatrix} \\ &= L \end{aligned} \qquad (36)$$

where the last equality follows from the well-known properties of the binomial distribution and where the approximation is virtually exact for large L, say $L \geq 32$. Strictly speaking, we can say only that (36) holds when each user is sending the same encoded bit in successive symbol periods, but the essential randomness provided by the m-sequence ensures that (36) still holds for any transmitted bit patterns from users $\mathbf{x}$ and $\mathbf{y}$.

Welch's bound (16) for the worst-case variance σ^2_{wc} in an S-CDMA code sequence set is actually a bound on the **average variance** σ^2_{avg} **over all users** and, on a per-user interference basis, is

$$\tilde{\sigma}^2_{wc} \geq \tilde{\sigma}^2_{avg} \geq L\frac{M-L}{M-1}. \tag{37}$$

The conclusions from (36) and (37) are that (i) the *spreading-by-superposition scheme is virtually optimum for S-CDMA if the number of users is large ($M \gg L$)*, which is the case of interest for the QUALCOMM system, but that (ii) a conventional CDMA sequence set may give significantly better performance when the number M of users is on the order of L or less. In fact, as shown in section 2.4, the preferentially-phased Gold sequence set of $M = L+1$ sequences gives

$$\tilde{\sigma}^2_{wc} = \tilde{\sigma}^2_{avg} = 1 \tag{38}$$

which corresponds to equality in (37). [Appendix A gives additional sequence sets satisfying, or nearly satisfying, (38).] Thus, by comparison of (36) and (38), we see that *spreading-by-superposition is substantially non-optimum for S-CDMA compared to conventional CDMA spreading with the preferentially-phased Gold sequence set when $M = L+1$ or less.*

For both QS-CDMA and A-CDMA, (36) still gives a good approximation to the per-user inter-user interference when **spreading-by-superposition** is used, i.e., the system *is rather insensitive to the chip offset vector δ* as defined in Section 2.1. Because this inter-user interference is rendered virtually independent from symbol-to-symbol by the essential randomness of the successive L-tuples that appear within an m-sequence of period $T = 2^L - 1$, *there is no need to distinguish between the cases of slowly-varying and rapidly-varying chip offsets*, as was necessary for ordinary CDMA (cf. Sections 2.2.1 and 2.2.2), although a stricter analysis shows that performance is slightly better for the rapidly-varying case.

It remains to compare the performance of the preferentially-phased Gold code (with $M = L+1$ sequences) to that of the spreading-by-superposition scheme, as given by (36), when QS-CDMA or A-CDMA is used.

Consider first QS-CDMA with $\Delta = 1$, where Δ is the specified maximum chip offset as defined in Section 2.1. For this case, it was shown in Section 2.4 that (on a per-user basis) for the preferentially-phased Gold sequence set

$$\tilde{\sigma}^2_{wc}(1) = \frac{L^2+L-1}{L} \approx L+1 \tag{39}$$

and that essentially this worst case was experienced by all $L+1$ users. From (39) and (36), we see that the *worst-case performance for the preferentially-phased Gold sequence set is almost exactly the same as the performance of the spreading-by-superposition scheme.* However, for "rapidly-varying chip offsets" [which may be a reasonable assumption when $\Delta = 1$, particularly in an environment when the set of active users is rapidly changing due to voice-activated transmission and the like, as the chip offsets are then the small fluctuations around δ=0 of the synchronization system that is designed to keep the users in symbol synchronization] that are uniformly distributed over 1 chip interval, (30) applies and gives for the preferentially-phased Gold signal set on a per-user basis

$$\tilde{\sigma}^2_{wc}(1) \approx \frac{1}{6}L, \tag{40}$$

which is a 7.8 **dB advantage** compared to the spreading-by-superposition scheme. The conclusion is that, for $\Delta \approx 1$, *the preferentially-phased Gold sequence set for $M = L+1$ users offers substantial advantages over the spreading-by-superposition scheme if the chip-offsets are rapidly-varying.*

For larger Δ, in particular for $\Delta = L - 1$ (A-CDMA), the clear superiority of the preferentially-phased Gold sequence set vanishes. Indeed the argument that was used to prove Lemma 3 shows that

$$\sum_{m \neq n} < \mathbf{x}^{(n)}, T^j \mathbf{x}^{(m)} >^2 = L^2 + L - < \mathbf{x}^{(n)}, T^j \mathbf{x}^{(n)} >^2 \tag{41}$$

still holds for all n and for every $j, 1 \leq j \leq L-1$, when the preferentially-phased Gold sequence set is used, and hence that the left side of (41) is still lower-bounded by $L^2 + L - 1$ for all n. It is tempting to conclude that (39) still holds but this is false for two reasons. First, a fixed offset j (where $1 \leq j \leq \Delta$) does not simultaneously maximize all the terms on the left of (41) except when $\Delta = 1$. Second, the use of (41) assumes that the **even** crosscorrelation function for the two users determines the inter-user interference, which is true only when Δ is small (see section 2.1) or when the users are sending the same bits in successive symbol periods.

Considering the worst-case offset problem and assuming that Δ is small so that the even cross-correlation functions essentially determine the inter-user interference (see section 2.1), one can say that σ^2_{wc} is surely upper-bounded by the **sum** of the terms on the left of (41) for all j between 1 and Δ and hence that, on a per-user basis,

$$\tilde{\sigma}^2_{wc}(\Delta) \leq \Delta \frac{L^2 + L - 1}{L} \approx \Delta L \tag{42}$$

holds for the preferentially-phased Gold sequence set. Thus, *for small Δ (say $\Delta \approx 2$ or less), the worst-case performance of the preferentially-phased Gold sequence set will be comparable to the performance of the spreading-by-superposition scheme — if the phase offsets are rapidly varying then the averaging effect will give an advantage for the preferentially-phased Gold code set.*

It is only for large Δ, in particular for A-CDMA where $\Delta = L - 1$, that the spreading-by-superposition scheme can win out over the preferentially-phased Gold sequence set. This is suggested by (42) which, however, is a rather bad bound for large Δ and, as pointed out above, requires that the even crosscorrelation functions determine the inter-user interference. Preliminary investigations at ESTEC indicate that the odd crosscorrelation functions for the preferentially-phased Gold sequence set are comparably as good in general as the even crosscorrelation functions. If we tentatively make the hypothesis that these odd crosscorrelation functions are in fact as good as the even ones, then from the fact that the worst even crosscorrelation value in the Gold sequence set (see table 2.4.1) is (to a very good approximation) $\sqrt{2L}$ or $\sqrt{4L}$ according as $g = \gcd(2e, m)$ equals 1 or 2, it follows that

$$\tilde{\sigma}^2_{wc}(\Delta) \leq \begin{array}{ll} 2L & \text{if } g = 1 \\ 4L & \text{if } g = 2 \end{array} \tag{43}$$

holds for the preferentially-phased Gold sequence set for all $\Delta, 1 \leq \Delta \leq L-1$. It follows further that *the preferentially-phased Gold sequence set (with $g = 1$) is at worst 3 dB inferior to the spreading-by-superposition scheme, even when A-CDMA is used if the odd crosscorrelation functions of the Gold set are as good as the even ones.* The improbability of the worst case (on the average only half of the users will have cross-correlations of $\sqrt{2L}$ with a given user, the others will have the ideal -1 correlation) suggests that, in most cases for large Δ, the preferentially-phased Gold sequence set would be only slightly inferior to the spreading-by-superposition scheme.

In summary, we may conclude

(1) *The spreading-by-superposition scheme is virtually optimum when* $M \gg L$ *and is the clear option of choice.*

(2) *For* $M \approx L$*, the preferentially-phased Gold sequence set [as well as some of the other sequence sets described in Appendix A] is virtually optimum for small* Δ *(say* $\Delta \leq 2$*) and is thus the clear option of choice for S-CDMA and QS-CDMA.*

(3) *For* $M \approx L$ *and A-CDMA, the spreading-by-superposition scheme and the preferentially-phased Gold sequence set give about the same performance if the odd crosscorrelation functions of this Gold sequence set are comparably as good as the even ones — this appears to be true but would have to be verified for the particular Gold code selected. If the even and odd crosscorrelation functions are indeed comparably good, then the choice of option in this case should be made on the basis of other factors such as the relative ease of implementation.*

4 References

[1] R. DeGaudenzi, R. Viola, "A Novel Code Division Multiple Access System for High Capacity Mobile Communication Satellites," *ESA Journal*, Vol. 13, 89/4, pp. 303 -327.

[2] R. DeGaudenzi, R. Viola, "Baseband and Network Control Subsystem for VSAT Terminals Utilizing SS-CDMA, Appendix 1, Draft Work Statement and Technical Specifications, 30 Jan. 1990.

[3] L. R. Welch, "Lower Bounds on the Maximum Cross Correlation of Signals," *IEEE Trans. Inform. Theory*, vol IT-20, pp. 397-399, May 1974.

[4] J. E. Mazo, " Some Theoretical Observations on Spread-Spectrum Communications," *The Bell System Technical Journal*, Vol. 58, No. 9, pp. 2013-2023, November 1979.

[5] D. V. Sarwate, M. B. Pursley, "Crosscorrelation Properties of Pseudorandom and Related Sequences," *Proceedings of the IEEE*, Vol. 68, No. 5, pp. 593-619, May 1980.

[6] V. M. Sidel'nikov, "On Mutual Correlation of Sequences," *Soviet Math. Dokl.*, Vol. 12, pp. 197-201, 1971.

[7] M. K. Simon, J. K. Omura, R. A. Scholtz, B. K. Levitt, *Spread Spectrum Communications,* Computer Science Press, 1985.

[8] R. J. McEliece, *Finite Fields for Computer Scientists and Engineers,* Kluver Intern. Series in Engineering and Computer Science, Vol. 23, 1987.

[9] W. W. Peterson, E. J. Weldon, Jr., *Error-Correcting Codes,* MIT Press, 1988.

[10] F. J. MacWilliams, N. J. A. Sloane, *The Theory of Error-Correcting Codes,* North-Holland Mathematical Library, Vol. 16, 1977.

[11] T. Verhoeff, "An Updated Table of Minimum-Distance Bounds for Binary Linear Codes," *IEEE Trans. on Inform. Theory*, Vol. IT-33, No. 5, September 1987.

[12] A. J. Viterbi, "Very Low Rate Convolutional Codes for Maximum Theoretical Performance of Spread-Spectrum Multiple-Access Channels," *IEEE J. Selected Areas in Comm.*, Vol. 8, No. 4, pp. 641-649, May 1990.

[13] A. J. Viterbi, "Spead Spectrum Communications — Myths and Realities," *IEEE Commun. Mag.*, Vol. 17, pp. 11-18, May 1979.

[14] J. L. Massey, "Uniform Codes," *IEEE Trans. on Inform. Theory*, Vol. IT-12, No. 2, pp. 132-134, April 1966.

[15] A. J. Viterbi and J. K. Omura, *Principles of Digital Communications and Coding.* New York: McGraw-Hill, 1979.

[16] G. S. Lauer, "Some Optimal Partial-Unit-Memory Codes," *IEEE Trans. on Inform. Theory*, Vol. IT-25, No. 2, pp. 240-243, March 1979.

Appendix A: Some Alternative CDMA Sequence Sets

Some code sequence sets in addition to the preferentially-phased Gold sequence set $\mathcal{G}$ considered in the body of this report are presented here. Except for the Hadamard code sequence set, they do not meet the ESTEC requirements on the crosscorrelation properties around the origin. However, all of these code sequence sets are almost optimal for S-CDMA with respect to Criterion 1. One of the main interests in the study of these alternative codes was to study the relation between the maximal crosscorrelation value c_{max} and the per-user interference parameter $\tilde{\sigma}_{wc}(\Delta)$ that essentially determines the performance of a QS-CDMA system. Note that $\tilde{\sigma}_{wc}(\Delta)$ depends only on the **even** crosscorrelation functions; thus for A-CDMA this parameter will generally give an optimistic approximation to the performance of the system because the influence of the **odd** crosscorrelation functions can strongly increase the interuser interference. Despite its importance, $\tilde{\sigma}_{wc}(\Delta)$ has been much less well studied than the maximal crosscorrelation c_{max}, and general results, beyond the trival fact that $\tilde{\sigma}_{wc}(\Delta) \leq c_{max}$, are scarce.

In case of S-CDMA, the per-user interference parameter $\tilde{\sigma}_{wc}(0)$ can be simply determined from the weight and distance distribution of the binary code corresponding to the code sequence set. In view of the examples below (see Table A.1), it turns out that for a code sequence set obtained from a "good" binary code, the Welch bound (21) is tight for $\tilde{\sigma}_{wc}(0)$ and, moreover, if the size M of the code sequence set is not too large, then the interference parameter $\tilde{\sigma}_{wc}(0)$ is much smaller than the maximal crosscorrelation value c_{max}.

In case of A-CDMA, little is known about the exact value of $\tilde{\sigma}_{wc}(L-1)$. In general, we can only lower bound $\tilde{\sigma}_{wc}(L-1)$ by the Welch bound (21), which does not seem to be tight. The evidence for the looseness of the Welch bound (21) in the case of A-CDMA comes from two examples, the sequence sets for the Gold code and the irreducible binary [85,8,40] BCH-code, where $\tilde{\sigma}_{wc}(L-1)$ coincides with c_{max}, which for the Gold code is about twice or four times the Welch lower bound depending on whether $L = 2^m - 1$ has odd m or even m respectively, and is about twice the Welch lower bound for the BCH-code.

Except for the sequence sets obtained from the Kerdock code and the two irreducible BCH-codes, the presented code sequence sets are well-known candidates for CDMA applications. We have considered the Kerdock code because it contains a large number of codewords; thus, it can be used to obtain a code sequence set for an S-CDMA system with a large number of possible users. On the other hand, the irreducible BCH-codes have been studied because they yield good sequence sets for the case where the number of users is only a small multiple of the length of the code sequences.

S-CDMA Code Sequence Sets:

1. The **Hadamard Code Sequence Set** (cf. e.g. Chap. 5.6 of [9]) consists of the $M = L = 2^m$ code sequences that form the rows of a ± 1 $2^m \times 2^m$ Hadamard matrix. Because the columns (as well as the rows) of this matrix are orthogonal, this code sequence set meets the Welch bound (21) with equality and thus is optimal for S-CDMA.

2. The **preferentially-phased Gold code** $\mathcal{G}$ (see Sec. 2.4) is optimal for S-CDMA, as was shown in Section 2.4 above.

3. **All phases of the small set of Kasami sequences**, which give a set of $M = L\sqrt{L+1}$ code sequences of length $L = 2^{2n}$. The small set of Kasami sequences have a very small maximal crosscorrelation value, $c_{max} = 1 + \sqrt{L+1}$ (cf. [5]). The Welch bound (21) yields $L - \sqrt{L+1} < \tilde{\sigma}^2_{wc}(0)$, thus, one has $-1 + \sqrt{L+1} < \tilde{\sigma}_{wc}(0)$. On the other hand, $\tilde{\sigma}_{wc}(0) \leq c_{max}$, and hence, one obtains the following good approximation: $\tilde{\sigma}_{wc}(0) \approx c_{max}$.

4. The **Kerdock code** is a nonlinear non-cyclic subcode of the 2nd order Reed-Muller code and it has well-known weight and distance distributions that coincide (cf. Fig. 15.7 of [10]). The Kerdock code contains $M = L^2$ codewords of length $L = 2^{2n}$. To every codeword $\mathbf{a}$ corresponds exactly one "antipodal" codeword $\mathbf{b}$, i.e., the Hamming distance $d(\mathbf{a}, \mathbf{b}) = L$ is maximal. The crosscorrelation between such antipodal codewords is $C_{\mathbf{ab}}(0) = L - 2d(\mathbf{a}, \mathbf{b}) = -L$. To avoid these large crosscorrelation values for S-CDMA, one takes half of the non-zero codewords to correspond to code sequences and one picks these in such a way that only one of the two antipodal pairs are in the code sequence set. The size of this code sequence set is thus $M = 2^{2n-1}$ and the weight (distance) distribution of the subcode used to determine the code sequences is given in the following table

i	A_i	c_i
$2^{2n-1} - 2^{n-1}$	$2^{2n-1}(2^{2n-1} - 1)$	2^n
2^{2n-1}	$2^{2n} - 1$	0
$2^{2n-1} + 2^{n-1}$	$2^{2n-1}(2^{2n-1} - 1$	-2^n

where i denotes the Hamming weight, A_i the number of codewords of Hamming weight i, and $c_i = L - 2i$ the corresponding correlation value. Summing up all correlation terms, one gets

$$\tilde{\sigma}^2_{wc}(0) = \frac{1}{M-1} \sum_i c_i^2 A_i \approx 2^{2n} - 2 = L - 2.$$

The Welch bound (21) yields almost the same value; hence, this code is almost optimal for S-CDMA. The maximal correlation value is $c_{max} = \sqrt{L}$.

5. Taking **all L phases of the sequences in the Gold sequence set**, one obtains an enlarged code sequence set with $M = L(L+2)$ codewords of length $L = 2^m - 1$ (cf. [5]). [The set of all Gold sequences contains $L + 2$ sequences of length L, exactly one sequence more than the preferentially-phased Gold sequence set $\mathcal{G}$.] The weight distribution of the Gold sequences is known and can be derived from Table 2.4.1. Proceeding as above for the Kerdock code, one finds $\tilde{\sigma}_{wc}(0) \approx \sqrt{L-1}$, which is almost same as the Welch bound (21). The maximal crosscorrelation value is $c_{max} = 1 + 2^{\lfloor (m+2)/2 \rfloor}$.

6. The **irreducible [85, 8, 40] BCH-code** is the best linear code of length $L = 85$ and dimension $K = 8$ (cf. [11]). Besides the zero-codeword, which is not used for a code sequence, the code consists of 3 different cycle sets all of period 85. The Hamming weights of the 3 cycle sets are 40, 40 and 48. Thus, the corresponding code sequence set has $M = 3 \times 85 = 255$, $c_{max} = 11$ and $\tilde{\sigma}_{wc}(0) = 7.53 \approx \sqrt{\frac{2}{3}L}$. Again, the Welch bound (21) is very tight.

7. The **irreducible [585, 12, 280] BCH-code** of length $L = 585$ consists, in addition to the zero-codeword that is not used for a code sequence, of 7 different cycle sets all of period 585. Three cycles have weight 280, one cycle has weight 296 and three cycles have weight 304. Thus, the corresponding code sequence set has $M = (3+1+3) \times 585 = 4095$. The correlation

parameters can readily be computed as $\tilde{\sigma}_{wc}(0) = 22.3 \approx \sqrt{\frac{6}{7}L}$ and $c_{max} = 25$. Again, the Welch bound (21) is very tight.

M	L	Code Sequence Set	$\tilde{\sigma}_{wc}(0)$	c_{max}
L	2^m	Hadamard code	0	0
$L+1$	2^m-1	Preferentially-phased Gold code $\mathcal{G}$	1	1
$L\sqrt{L+1}$	$2^{2n}-1$	All phases of small set of Kasami sequ.	$\approx \sqrt{L+1}$	$1+\sqrt{L+1}$
$\approx \frac{1}{2}L^2$	2^{2n}	Half of Kerdock code	$\approx \sqrt{L-2}$	$\sqrt{L}$
$L(L+2)$	2^m-1	All phases of sequences in Gold code	$\approx \sqrt{L-1}$	$1+2^{\lfloor (m+2)/2 \rfloor}$
$3L$	85	[85, 8, 40] BCH-code, non-zero codewords	7.53	11
$7L$	585	[585, 12, 280] BCH-code, non-zero codewords	22.3	25

Table A.1: Parameters of Some Alternative S-CDMA Code Sequence Sets.

A-CDMA Code Sequence Sets:

Only **even** crosscorrelation functions are considered. [The odd crosscorrelation functions are of the same importance for A-CDMA system performance but little is known about them.] The investigated A-CDMA code sequences are listed in Table A.2. The interference parameter $\tilde{\sigma}_{wc}(L-1)$ could only be determined for the Gold sequences and the sequences from the [85, 8, 40] BCH-code, where $\tilde{\sigma}_{wc}(L-1) = c_{max}$. For the other code sequence sets, we have given the lower bound (21) instead of the interference parameter $\tilde{\sigma}_{wc}(L-1)$. The Welch bound (21) is seen to be not very tight, except for the small set of Kasami sequences.

M	L	Code Sequence Set	$\tilde{\sigma}_{wc}(L-1)$	c_{max}
$\sqrt{L+1}$	$2^{2n}-1$	Small set of Kasami sequences	$> -\frac{3}{2}+\sqrt{L+1}$	$1+\sqrt{L+1}$
$L+2$	2^m-1	Gold sequences	$1+2^{\lfloor (m+2)/2 \rfloor}$	$1+2^{\lfloor (m+2)/2 \rfloor}$
$\approx L\sqrt{L+1}$	$2^{2n}-1$	Large set of Kasami sequences	$> \sqrt{L-\frac{2}{\sqrt{L+1}}}$	$1+2\sqrt{L+1}$
3	85	[85, 8, 40] BCH-code, non-zero codewords	11	11
7	585	[585, 12, 280] BCH-code, non-zero codewords	> 20.29	25

Table A.2: Parameters of Some Alternative A-CDMA Code Sequence Sets.

Appendix B: Proof of Independency of δ_{ij} and $\mathbf{Z}_{ij}$

Here we show the statistical independency of the relative chip offsets $Y_{ij} = \delta_{ij}$ and the random variable $Z_{ij} = cos(\omega\delta_{ij}T_c + \phi_{ij})$, the proof of which was deferred from Section 2.1 to this Appendix. Recall that the relative chip offsets $Y_{ij} = \delta_{ij} = \delta_j - \delta_i$ and the relative carrier-phase offsets $\phi_{ij} = \phi_j - \phi_i$ are statistically independent and that ϕ_{ij} is uniformly distributed in the interval from 0 to 2π. To prove the statistical independency of δ_{ij} and $Z_{ij} = cos(\omega\delta_{ij}T_c + \phi_{ij})$, we shall use the following useful criterion stated in the next lemma.

Lemma: If $E[f(X)g(Y)] = E[f(X)]E[g(Y)]$ for all continuous functions f and g for which these expectations exist, then X and Y are statistically independent, and conversely.

Proof: Choose f and g as continuous approximations to impulses located at x_0 and y_0, respectively, i.e.,

$$f(X) \approx \delta(X - x_0) \qquad\qquad g(Y) \approx \delta(Y - y_0).$$

Computing expectations, one obtains

$$\begin{aligned} E[f(X)g(Y)] &= p_{XY}(x_0, y_0) && \text{if } p_{XY} \text{ is continuous at } (x_0, y_0), \\ E[f(X)] &= p_X(x_0) && \text{if } p_X \text{ is continuous at } x_0, \\ E[g(Y)] &= p_Y(y_0) && \text{if } p_Y \text{ is continuous at } y_0. \end{aligned}$$

Using the assumption, one gets $p_{XY}(x_0, y_0) = p_X(x_0)p_Y(y_0)$ if p_{XY} is continuous at (x_0, y_0). Upon taking a continuous approximation $\tilde{p}_{XY}$ to the probability density function p_{XY}, the statistical independency of X and Y follows.

The converse part of the lemma is trivial.

Application: Statistical Independency of δ_{ij} and Z_{ij}

Take $X = Z_{ij}$ and $Y = \delta_{ij}$, and note that the conditional expectation

$$E[f(X)|Y = y] = E[f(cos(\omega\delta_{ij}T_c + \phi_{ij}))|Y = y] = E[f(cos(\omega yT_c + \phi_{ij}))|Y = y] = E[f(cos(\omega yT_c + \phi_{ij}))],$$

(where at the last step we have used the independency of Y and ϕ_{ij}) does not depend on y because ϕ_{ij} is uniformly distributed between 0 and 2π. The assumption of the lemma is now easily verified as follows:

$$\begin{aligned} E[f(X)g(Y)] &= \int_{-\infty}^{\infty} E[f(X)g(Y)|Y = y]p_Y(y)dy \\ &= \int_{-\infty}^{\infty} E[f(X)g(y)|Y = y]p_Y(y)dy \\ &= \int_{-\infty}^{\infty} g(y)E[f(X)|Y = y]p_Y(y)dy \\ &= E[f(cos(\omega yT_c + \phi_{ij}))]\int_{-\infty}^{\infty} g(y)p_Y(y)dy \\ &= E[f(X)]E[g(Y)]. \end{aligned}$$

The statistical independency of δ_{ij} and Z_{ij} now follows from the lemma.