Towards a Cryptanalysis of Spectral-Phase Encoded Optical CDMA with Phase-Scrambling

Sharon Goldberg

September 30, 2006

Abstract

We show how an eavesdropper with some specific knowledge of the traffic sent over a spectral-phase encoded optical CDMA system with phase-scrambling can break the confidentially of certain systems within a few bit intervals.

Contents

1 Introduction

Optical CDMA (OCDMA) [1] is a particularly attractive alternative to traditional digital encryption because it has the potential to perform encryption at ultra-high data rates by initializing passive optical components (e.g. phase masks, delay lines) according to a secret key that needs only occasional updates. To be a viable alternative to digital encryption, OCDMA systems should maintain data confidentiality even when these optical components are reconfigured (*i. e.* the key is refreshed) at rates much slower than aggregate system data rates. We focus on spectral-phase-encoded OCDMA systems using phase-scrambling [2, 3], which have emerged as the leading proposal for providing data confidentiality at the physical layer. Typical OCDMA schemes require some orthogonality between codewords and are therefore restricted to use codes with low cardinality [4, 5]. These schemes are therefore vulnerable to brute force searches by an eavesdropper who cycles through all possible codewords in an effort to find one that results an ungarbled datastream. On the other hand, a spectral-phase scrambling scheme offers a keyspace that grows *exponentially* with the number of frequency bins used, so that brute-force searches can be made infeasible. We assume that the *secret key* used in the system is the setting of the phase-scrambler, and analyze this system using the assumptions of cryptanalysis [6]. In particular, we explore *known plaintexts attacks* in which an eavesdropper obtains the encryption of some set of known messages, and uses this information to learn the secret key. Our first contribution is to show circumstances in which confidentiality is determined by the parallelism (*i. e.* the number of users) in the system, rather than by the number of frequency bins used for encoding. Our next contribution is to show that even when some systems are highly parallelized (*i. e.* have a large number of users), an eavesdropper can still learn the key with high probability after only two bit intervals. Our results thus far suggest that to maintain confidentiality when the secret key is the phase-scrambler setting, components should be tuned at rates comparable to the system data rates.

1.1 System Overview

In spectral-phase-encoded OCDMA, pulse streams are encoded by adjusting the phase of their frequency components using all-optical pulse shaping techniques. For each optical pulse, the encoding process consists of dividing the pulse spectrum into W frequency bins, applying one of C possible discrete phase shifts at each frequency (as prescribed by the choice of codeword), and recombining the frequency bins to produce the coded pulse. Decoding is achieved in a similar manner, by applying inverse phase-shifts at

each frequency bin [1]. We envision a scheme where a group of N time-synchronized OCDMA users (*i. e.* transmitter-receiver pairs) simultaneously share a waveband of a single optical fiber (that may then be overlayed onto a WDM network, or shared by another group OCDMA of users). Each user's encoded pulse train is passively coupled onto the waveband, transmitted along the fiber, and then the superposed encoded pulse trains are passively split and individual pulse streams are recovered at each decoder. To prevent energy detection attacks on on/off keyed OCDMA [5], we consider 2-code keyed (2CK) systems where each user uses a 'bright' codeword to send a '1' bit, and a different 'dark' codeword to send a '0' bit. (Other constant-energy modulation formats such as phase-shift keying are also possible.) The field of the optical signal sent by user j is represented as the signal

$$\alpha_j(t) = \sum_{i=1}^{W} \cos(f_i t + \theta_{i,j}(t)) \tag{1}$$

where f_i the the frequency of the i^{th} wavelength in the OCDMA waveband (i.e. $f_i = \frac{c}{\lambda_i}$ where c is the speed of light), and $\theta_{i,j}(t)$ is the phase shift applied on the i^{th} wavelength according to the particular codeword sent by user j during the bit interval containing time t. The n different codestreams aggregated onto a single fiber using passive coupling. Because the different lengths of the fibers between each user j's transmitter and the passive coupler, a relative phase shift $\phi_j(t)$ that depends of laser intensity fluctuations and temperature fluctuations will be introduced into user j's codestream. We will model this phase shift as a random variable on $[0, 2\pi]$. Then, aggregation of the N codestreams is equivalent to summation of each codestream. We can express the field of optical signal composed of the N aggregated codestreams as

$$\rho(t) = \sum_{j=1}^{N} \sum_{i=1}^{W} \cos(f_i t + \theta_{i,j}(t) + \phi_j(t)) \tag{2}$$

Spectral phase OCDMA systems use orthogonal code families to reduce MAI [1], so that number of available codewords is typically limited to W, making the system vulnerable to brute force attacks. To eliminate this vulnerability, an additional scrambling (descrambling) encoder that can take on any of C^W possible spectral-phase settings, is placed immediately after the coupler (before the splitter) as in Fig. **??**. We assume that the setting of the shared scrambler is the secret key that is used to prevent an eavesdropper from understanding the messages sent along the OCDMA waveband [2, 3]. Since there are now C^W (as opposed to just W) possible keys, when W is large (say $W \approx 70$) a simple

brute force search through the keyspace is no longer possible. Finally, to avoid attacks on a link carrying only a single user's encoded pulse stream [5, 7], links between each transmitter and the passive coupler are physically secured so that an eavesdropper can only tap the line between the scrambler and descrambler. We can expresses the field of the scrambled superposition of N encoded codestreams at the eavesdropper's tap as

$$\chi(t) = \sum_{j=1}^{N} \sum_{i=1}^{W} \cos(f_i t + \theta_{i,j}(t) + \phi_j(t) + k_i) \tag{3}$$

where f_i is the the frequency of the i^{th} frequency bin, and $\theta_{i,j}(t)$ is the phase shift applied on the i^{th} frequency bin according to the particular codeword sent by user j during the bit interval containing time t, k_i is the phase setting of the i^{th} frequency bin in the scrambler (which we assume remains the same for the duration of the eavesdropper's attack), and ϕ_j is the *randomly time-varying* phase shift of user j's encoded pulse stream relative to all other other users' encoded pulse streams caused by variations in the lengths of the fibers between user j's transmitter and the passive coupler, and time-varying fluctuations in laser intensity or temperature.

1.2 Background: Encryption Schemes

An encryption scheme uses a secret key k to encrypt a plaintext message to a ciphertext. When analyzing an encryption scheme, it is standard practice to apply Kerckhoffs' Principle [6], which states that that a cryptosystem should be secure even if everything about the system, except from the secret key, is *public knowledge*. In a well-designed system, only the key needs to be secret; in fact, when cryptanalyzing the system, everything apart from the secret key should be assumed to be public.

We now enumerate some standard cryptanalytic attacks on encryption schemes, ranked in order of increasing difficulty for the eavesdropper:

1. *Ciphertext Only Attacks (COA):* The eavesdropper obtains a set of ciphertexts, and uses these to learn the secret key.

2. *Known Plaintext Attacks (KPA):* The eavesdropper *knows* some set of plaintexts and obtains the corresponding ciphertexts, and uses these to learn the secret key.

3. *Chosen Plaintext Attacks (CPA):* The eavesdropper has the capability to *choose* plaintexts to be encrypted and obtains the corresponding ciphertexts, and uses these to learn the secret key.

(Note that once the eavesdropper learns the key, she has completely broken the system since she is able to decrypt all ciphertexts.) Since the COA-attack is the easiest attack for the eavesdropper to perform, a scheme that is secure against COA-attacks (*i. e.* COA-secure) is less secure than a scheme that is KPA-secure. Thus, while COA-security is the weakest form of security, to date it is the only form of security considered in the literature on OCDMA [3, 4, 5, 7]. (Note the *minimal* threshold for security of standard digital encryption schemes is the ability to withstand CPA-attacks [8] §3.2.1.) In this paper, we focus on KPA-attacks. The KPA-attack is a realistic threat, since data traffic is never completely random; it contains packet headers, 'hello' packets, or framing elements (*e.g.* SONET framing) that are publicly known and may be used to launch a KPA-attack.

1.3 Security Parameters of the System

We now clearly define the plaintext, secret key, and ciphertext in our system:

- **Plaintext:** In this paper, we assume that for the system in Fig. **??**, the set of 'bright' and 'dark' codewords assigned to each OCDMA user is *not* part of the secret key. (Note that schemes in which codewords assigned to each user are kept secret are also possible [2].) Therefore, by Kerckhoffs' principle, we assume that this information is known to the eavesdropper. Then, for each bit-interval, instead of defining the *plaintext* as the bits transmitted by each of the N users during that bit-interval, we define the plaintext as the set of *codewords* transmitted by each of the N users during that bit-interval. (Recall that encoders and scramblers are restricted to use phase shifts of $0, \frac{2\pi}{C}, ..., \frac{2\pi(C-1)}{C}$.) We represent the plaintext as an $N \times W$ $\{0, \frac{2\pi}{C}, ..., \frac{2\pi(C-1)}{C}\}$-matrix Θ, where entry $\theta_{i,j}$ gives the phase shift applied to frequency bin i corresponding to the codeword send by user j.

- **Secret Key:** The key is the set of W phase shifts $k_i \in \{0, \frac{2\pi}{C}, ..., \frac{2\pi(C-1)}{C}\}$ applied by the scrambler at frequency bin i. That is, the secret key is a vector $\mathrm{k} \in \{0, \frac{2\pi}{C}, ..., \frac{2\pi(C-1)}{C}\}^W$.

- **Ciphertext:** The ciphertext is $\chi(t)$ in (3), the optical signal seen at the eavesdropper's tap. We assume that the eavesdropper detects the ciphertext $\chi(t)$ using optical beat detection [5, 9] by passing $\chi(t)$ through a WDM demultiplexer to obtain $\chi_i(t)$ for $i = 1, ..., W$ (where $\chi_i(t)$ is $\chi(t)$ filtered at i^{th} frequency bin). By interfering $\chi_i(t)$ with a local oscillator signal $\cos(f_i t)$, then detecting $\chi_i(t) + \cos(f_i t)$, with

a photodetector (which operates a square law device with an envelope detector or lowpass filter) to produce a signal

$$\begin{aligned} y_i(t) &= LPF((\chi_i(t) + \cos(f_i t))^2) \\ &= LPF((\textstyle\sum_{j=1}^{N} \cos(f_i t + \theta_{i,j}(t) + \phi_j(t) + k_i) + \cos(f_i t))^2) \\ &= LPF(2 \textstyle\sum_{j=1}^{N} \cos(f_i t + \theta_{i,j}(t) + \phi_j(t) + k_i) \cos(f_i t)) \\ &= LPF(\textstyle\sum_{j=1}^{N} (\cos(2 f_i t + \theta_{i,j}(t) + \phi_j(t) + k_i) + \cos(\theta_{i,j}(t) + \phi_j(t) + k_i))) \\ &= \textstyle\sum_{j=1}^{N} \cos(\theta_{i,j}(t) + \phi_j(t) + k_i) \end{aligned} \tag{4}$$

The eavesdropper can obtain signals $y_i(t)$ for each wavelength $i = 1, 2, ..., W$. We will call the vector $\mathrm{y}(t_0) = [y_1(t_0)\, y_2(t_0) \ldots y_W(t_0)]^T$ the *adversary's measurement at time* t_0 or more concisely, *the measurement.*

2 Size of Search Space in a KPA attack

We begin by determining the size of the keyspace of the system. Recall that the key $\mathrm{k} \in_R \{0, \frac{2\pi}{C}, ..., \frac{2\pi(C-1)}{C}\}^W$. Therefore the number of possible keys is C^W [1]. The our first new contribution is to show the size of the keyspace is actually much smaller than C^W. Recall that for a coherent, orthogonal spreading code such as the Hadamard code, the number of codes available is equal to the number of code elements. Therefore, we know that $N \leq W$. Furthermore, if 2-code-keying is used, it follows than $N \leq \frac{W}{2}$.

Claim 2.1 *In the KPA-setting, the size of the space in an exhaustive search for* all W *elements of the key is* C^N.

Proof: Consider an eavesdropper that a obtains a measurement y of all W wavelengths of the ciphertext (simultaneously) at some time t_0 using beat detection as in (4). Fixing time at t_0, and therefore dropping the time index from (4), we can write the measurements y as W beat detection equations:

$$\begin{aligned} y_1 &= \textstyle\sum_{j=0}^{n} \cos(\phi_j + \theta_{1j} + k_1) \\ y_2 &= \textstyle\sum_{j=0}^{n} \cos(\phi_j + \theta_{2j} + k_2) \\ ... &= ... \\ y_w &= \textstyle\sum_{j=0}^{n} \cos(\phi_j + \theta_{wj} + k_w) \end{aligned} \tag{5}$$

[1] [2, 3] have shown that the strongest known attack on a system with $N > 1$ users is to have the eavesdropper do a brute force search through a space of size C^W.

The system of W equations (5) has N unknowns on $[0, 2\pi]$, and W unknowns on $\{0, \frac{2\pi}{c}, ..., (c-1)\frac{2\pi}{c}\}$. Thus, since the eavesdropper knows the θ_{ij} in a known plaintext attack, if the eavesdropper guesses the first N elements of k (that is, if he guesses $(k_1, k_2, ..., k_N)$), he can use solve the first N beat detection equations to obtain a guess for $\phi = [\phi_1\, \phi_2 \ldots \phi_N]^T$. He can then use his guess of ϕ to solve the last $W - N$ beat detection equations for $(k_{N+1}, ..., k_W)$. Because there are C^N possible values for the first n elements of k, the true size of the keyspace in a known-plaintext-attack is C^N. ∎

Claim 2.1 shows that confidentiality is determined by the the amount of parallelism in the system, (the number of parallel OCDMA users n), rather than by the number of frequency bins W used for in encoding. This can be significant reduction in the key search space, since systems are typically designed so that $N < W$ (*e.g.* in 2-code-keying systems using orthogonal codes $N \leq \frac{W}{2}$). As an example, in the KPA-setting, a system using a large number of frequency bins $W = 70$ but only a small number of users $N = 4$ has a key search space of size only $2^4 = 16$ rather than 2^{70}, as originally mentioned in [2, 3].

3 KPA Attack with 2 Known Plaintexts

Even when the system is highly parallelized so that N is large enough to prevent brute force attacks (say $N \approx 70$ users), we now present another KPA-attack that eavesdropper can use reduce the size of the key search space from C^N to an even smaller set of possibilities.

Claim 3.1 *Suppose the eavesdropper in a 2-code-keying system obtains* two known plaintexts *and corresponding measurements* $(\Theta_1, \mathrm{y}_1), (\Theta_2, \mathrm{y}_2)$. *Then the eavesdropper has* $2W$ *equations and* $W + 2N$ *unknowns, that can be solved for the* W *elements of the key (when* $N \leq \frac{W}{2}$*). We shall show that the eavesdropper's key search space is reduced from* C^N *to the set of solutions to this system of* $2W$ *equations. Furthermore, if there is a unique solution to these* $2W$ *equations then the eavesdropper immediately learns the key.*

In this section, we first provide a mathematical framework for the problem. We then show how the eavesdropper reduce his key search space in the KPA from C^N to the smaller set of solutions. Finally, we present case study of the success of this attack on a 2-code-keying OCDMA system using binary phase shifts and the Hadamard codes.

3.1 A Mathematical Framework

We now express (5) in matrix format. Start by defining a vector x of the cosine of the inter-user phases ϕ_j as

$$\mathrm{x} = \begin{bmatrix} \cos\phi_1 \\ \cos\phi_2 \\ \dots \\ \cos\phi_N \end{bmatrix}$$

Next, we map the key elements and θ_{ij} from radians to complex numbers using the mapping $\gamma \rightarrow e^{i\gamma}$ where i is the imaginary number. (For example, for binary phase shifts $\theta_{ij} \in \{0, \pi\}$ the mapping from radians to integers is $0 \rightarrow 1, \pi \rightarrow -1$, so that we can write $\theta_{ij} \in \{1, -1\}$.) We can now write (5) as

$$\begin{array}{rcl} y_1 & = & k_1 \sum_{j=0}^{N} \theta_{1j} \cdot \cos\phi_j \\ y_2 & = & k_2 \sum_{j=0}^{N} \theta_{2j} \cdot \cos\phi_j \\ \dots & = & \dots \\ y_{2n} & = & k_W \sum_{j=0}^{N} \theta_{Wj} \cdot \cos\phi_j \end{array} \tag{6}$$

and in matrix form this can be written as

$$\mathrm{y} = K\Theta^T\mathrm{x} \tag{7}$$

where $\mathrm{y} = [y_1 y_2 ... y_W]^T$ is a vector of measurements, and $K = diag([k_1 k_2 ... k_W])$ is a $W \times W$ diagonal matrix of the key elements where entry k_i is a complex number representing the phase shift applied by the scrambler at frequency bin i, and Θ is a $N \times W$ matrix, where entry $\theta_{i,j}$ is a complex number representing the phase shift applied to frequency bin i corresponding to the codeword send by user j.

3.2 Details of the Attack

Consider an eavesdropper with a two measurements of the ciphertext $\mathrm{y}_1, \mathrm{y}_2$ obtained using beat detection corresponding to two *known* plaintexts Θ_1 and Θ_2. We now show how this eavesdropper in the KPA setting can cut down the size of the key search space from C^N to a very small set of possibilities for the key. For the purpose of this analysis, we will assume that the eavesdropper makes perfect measurements of $\mathrm{y}_1, \mathrm{y}_2$, free from noise. In future papers we will extend this analysis to noisy measurements.

Note that we will use the notation $K = diag(\mathrm{k})$, and $Y_a = diag(\mathrm{y})$ in the rest of this

report to switch between representations of diagonal matrices (e.g. K) and vectors (e.g. k). Now, since Y and K are diagonal matrices, for each measurement we can rewrite (7) as

$$Y\mathrm{k} = \Theta^T \mathrm{x} \tag{8}$$

Since the codewords sent by different users are orthogonal, it follows that Θ^T has a least N linearly independent columns, so the columns of Θ^T can be partitioned into an *invertible* $N \times N$ matrix A^T, and an $(W-N)\times N$ matrix B^T. We use the same partition to partition Y into Y_a, Y_b and K into K_a, K_b. We can now write (8) into a set of equations

$$\begin{aligned} Y_a\mathrm{k}_a &= A^T\mathrm{x} \\ Y_b\mathrm{k}_b &= B^T\mathrm{x} \end{aligned} \tag{9}$$

Now, using the fact the the fact that A is invertible, we can solve these two matrix equations for x to find that

$$Y_b^{-1}B^T{A^T}^{-1}Y_a\mathrm{k}_a = \mathrm{k}_b \tag{10}$$

Now suppose the two plaintext-measurement pairs (Θ_1, y_1), (Θ_2, y_2) are such that Θ_1 and Θ_2 can be divided along the *same partition* to form two invertible matrices A_1 and A_2. (We show in §3.3.1 that this is always be the case in a 2-code-keyed OCDMA system using the Hadamard codes with $W = 2N$). Then since y, x and possibly Θ do change between measurements, but k does not, it follows that the eavesdropper has two matrix relations of the form (10) that can be equated to find obtain the simple relation

$$Q\mathrm{k}_a = 0$$

where

$$Q = Y_{b1}^{-1}B_1^T{A_1^T}^{-1}Y_{a1} - Y_{b2}^{-1}B_2^T{A_2^T}^{-1}Y_{a2} \tag{11}$$

Therefore the problem of learning k reduces to the problem of finding all valid $\widehat{\mathrm{k}}_a$ that satisfy $Q\widehat{\mathrm{k}}_a = 0$, and then using (10) to obtain $\widehat{\mathrm{k}}_b$. It follows that the true key k is contained this set of $(\widehat{\mathrm{k}}_a, \widehat{\mathrm{k}}_b)$.

For ease of exposition, we will now assume that the scrambler is restricted to use only binary phase shifts $\mathrm{k} \in \{1, -1\}^W$, (so that for a guess $\widehat{\mathrm{k}}_a$ to be valid it must be in $\{1, -1\}^N$). However, our discussion applies to schemes with an arbitrary discrete phase shifts of size C. Suppose that there are exactly m linearly independent $\{1, -1\}^N$-vectors satisfying $Q\widehat{\mathrm{k}}_a = 0$. Recall that given Q, it is trivial to find these m vectors using

Gaussian elimination. By taking linear combinations of these m vectors using coefficients in $\{0, 1, -1\}$, it follows that the eavesdropper has less than 3^m possible valid guesses for the true key k_a. (More generally, with an arbitrary discrete phase shifts of size C, the eavesdropper has less than $(C+1)^M$ possible valid guesses for the true key k_a.)

Now, observe that if $m = 1$, then this is only one valid key guess $\widehat{\mathrm{k}}_a$ solving the equation $Q\widehat{\mathrm{k}}_a = 0$ so that eavesdropper has immediately learned the key. One the other hand, the skeptical reader might claim that if $m \approx n$ the eavesdropper gains nothing by solving $Q\widehat{\mathrm{k}}_a = 0$. However, the following case study shows that there is a high probability that M will be very small.

3.3 Success of the Attack: A Case Study

To quantify the size of the key search space after the attack with 2 known plaintexts, we have done a detailed analysis in of a N-user system using 2-code-keying with $W = 2N$ frequency bins and the standard $2N$-Hadamard codes, where key and codewords elements can take on binary phases ($C = 2$).

In order to make our analysis more concrete, we will assume that the system uses the standard $2N$-Hadamard code, obtained from the standard $2N$-Hadamard matrix derived recursively from

$$H_{2N} = \begin{bmatrix} H_N & H_N \\ H_N & -H_N \end{bmatrix} \quad H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{12}$$

Each codeword is a row of the H_{2N} matrix. Note also that H_N is a symmetric matrix. For ease of exposition, we assume that codewords are assigned as follows: for $j = 1...N$, user j sends a '0' using codeword j in the H_{2N} matrix, and sends a '1' using codeword $j + N$ in the H_{2N} matrix.[2]

3.3.1 Some Mathematical Preliminaries

We now discuss the structure of the $N \times 2N$ plaintext matrix Θ. Construct the plaintext matrix Θ so that the j^{th} row of Θ is the codeword sent by user j. Then, Claim 3.2 shows how any plaintext Θ may be partitioned into an invertible matrix A and another matrix

[2]Note that if codewords are assigned from the H_{2N} matrix according to a different scheme OR if a different instantiation of the H_{2N} matrix was used (e.g. $H' = QH_{2N}P$ where Q, P are $\{0, \pm 1\}$-permutation matrices) our analysis would be identical apart from the fact that a different partition of Θ into A, B would be required (instead of $\Theta = [A, B]$ with $A = H_n$, $B = LH_n$ as described in §3.3.1). Futhermore, as long as $W = 2N$, this partition could apply to all plaintexts Θ, since all plaintexts Θ would have the same set of N linearly independent columns.

B (as prescribed in §3.2).

Claim 3.2 *If user j sends a '0' bit using codeword j in the standard H_{2N} matrix, and sends a '1' bit using codeword $j+N$ in the standard H_{2N} matrix, then we can write any plaintext matrix Θ in block matrix form as*

$$\Theta = [H_N \quad LH_N]$$

where H_N is the standard Hadamard matrix and L is a diagonal matrix with ± 1 entries and $L_{jj} = +1$ when user j sent a '0' bit, and $L_{jj} = -1$ when user j sent a '1' bit. Therefore, the L matrix completely specifies the Θ matrix.

Proof: Observe that the j^{th} row of Θ is either two copies of j^{th} a row of H_N, or one copy of the j^{th} row of H_N followed by one copy of the j^{th} row of $-H_N$. Therefore we can write $\Theta = [A\, B]$ where $A = H_N$ and is invertible. Furthermore, we have that every row of B is equal to $\pm$ a row of H_N. Thus, $B = LH_N$ for L a diagonal matrix with ± 1 entries. Furthermore, when $L_{jj} = 1$, the j^{th} row of the Θ matrix is two copies of j^{th} a row of H_N. Equivalently, the j^{th} row of Θ is equal to the j^{th} row of the H_{2N} matrix, which, by to our assignment of codewords to users, implies that user j sent a 0 bit. Similar logic shows that $L_{jj} = -1$ when user j sent a '1' bit. ∎

Note that Claim 3.2 shows that all plaintexts can be partitioned along the same partition, namely $\Theta = [A\ B]$, as is required by §3.2. Now, consider an eavesdropper with two known plaintexts $\Theta_1 = [H_N\ L_1 H_N]$, $\Theta_2 = [H_N\ L_2 H_N]$ and two corresponding measurements $\mathrm{y}_1 = [\mathrm{y}_{a1}\ \mathrm{y}_{b1}]^T$, $\mathrm{y}_2 = [\mathrm{y}_{a2}\ \mathrm{y}_{b2}]^T$. Following the argument in §3.2, the eavesdropper obtains a set of key guesses $\widehat{\mathrm{k}} = [\widehat{\mathrm{k}}_a\, \widehat{\mathrm{k}}_b]^T \in \{1, -1\}^{2N}$, by solving for $\widehat{\mathrm{k}}_a$ in $Q\widehat{\mathrm{k}}_a = 0$, and using (10) to obtain $\widehat{\mathrm{k}}_b$. Observe that we can rewrite (11) using Claim 3.2 as

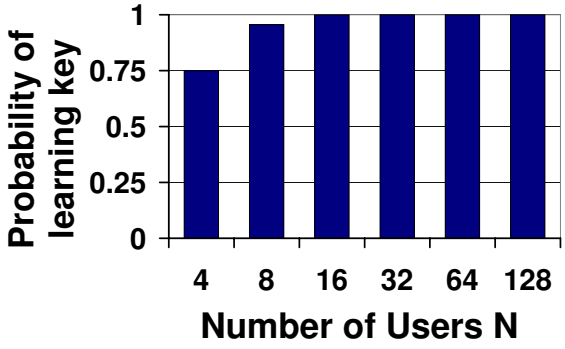
$$Q = Y_{b1}^{-1} H_N L_1 H_N Y_{a1} - Y_{b2}^{-1} H_N L_2 H_N Y_{a2} \tag{13}$$

3.3.2 Characterizing the eavesdropper's search space

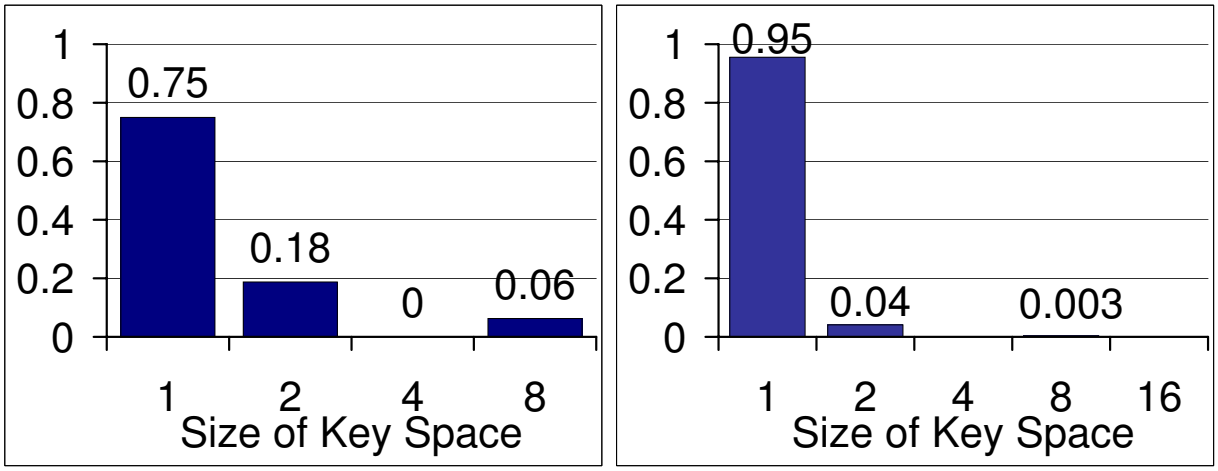
Recall that m is the number linearly independent $\{1, -1\}^N$-vectors satisfying $Q\widehat{\mathrm{k}}_a = 0$ with Q as in (13). Equivalently, m is the rank of the nullspace of Q, $m = N - rank(Q)$. Recall that m quantifies the size of the eavesdroppers key search space after this KPA attack, since the true key k_a is some function of these m vectors. In the balance of this section, we will characterize the size of the eavesdroppers key space after this KPA-attack by using m as our metric.

Claim 3.3 *Suppose the eavesdropper obtains any two arbitrary known plaintexts* Θ_1, Θ_1 *and corresponding noise-free measurements* $\mathrm{y}_1, \mathrm{y}_2$*. Then as the number of users* N *becomes large, the probability that* $m = 1$ *(so that the eavesdropper immediately learns the key* k *by solving* $Q\widehat{\mathrm{k}}_a = 0$*) approaches 1.*

Proof: We have not proved this theorem analytically. (We present an analytic proof of a slightly weaker version of this theorem in Claim 3.4). However, using simulation in MATLAB we found the relationship between number of users N and probability, over all possible known plaintexts pairs, of eavesdropper learning key (because of existence of unique solution to $Q\widehat{\mathrm{k}}_a = 0$, *i.e.* because $m = 1$). The probabilities for $N = 4, 8$ are exact. For large $N > 8$, exact enumeration became infeasible, so the probabilities were obtained from 10000 randomly sampled combinations of $(\mathrm{k}, \Theta_1, \Theta_2, \mathrm{x}_1, \mathrm{x}_2)$.



We now show the distribution of the size of the key search space after this KPA-attack, over all possible plaintext pairs, when N is small (i.e. $N = 4$, $N = 8$). Since Lemma A.3 shows that m and the size of the key search space is completely determined by Θ_1, Θ_2 (and is independent of k_a, k_b, x_1, x_2), we obtained these results for fixed k_a, k_b, x_1, x_2 over all 2^{2N} possible plaintext combinations for (Θ_1, Θ_2).



These distributions clearly indicate that $m = 1$ with high probability, and further

that as N increase, it is more probable that a randomly chosen pair (Θ_1, Θ_2) will result in a Q matrix such that $m = 1$. That is, as the security parameter N gets larger, an eavesdropper is more likely to break the security of the scheme if he obtains two arbitrary known plaintexts and corresponding measurement pairs. ∎

Since we have not been able to prove Claim 3.3 analytically, we prove analytically a claim that shows that for 75% of all possible known plaintext pairs, the adversary can learn the key.

Claim 3.4 *Suppose the eavesdropper obtains any two arbitrary known plaintexts* Θ_1, Θ_1 *and corresponding noise-free measurements* $\mathrm{y}_1, \mathrm{y}_2$. *Let* Θ_1, Θ_2 *be specified by diagonal* ± 1 *matrices* L_1, L_2 *as in Claim 3.2. Then if* any *of the following three matrices,* $L1$, L_2 *and* $L_1 L_2$, *contain an odd number of '-1's then the eavesdropper immediately learns the key* k *by solving* $Q\widehat{\mathrm{k}}_a = 0$ *(i. e.* $m = 1$*).*

Proof: We prove this in the Appendix. Note that the number of known plaintexts pairs that do break the system is $1 - (\frac{1}{2})^2 = 75\%$.

Taken together, our results indicate that for this $2N$-Walsh-Hadamard 2-code-keyed system, an eavesdropper with an arbitrary pair of known plaintexts has a very high probability of learning the key.

4 Future Work

In future papers we will study how noise in measurements $\mathrm{y}_1, \mathrm{y}_2$ affects our KPA-attacks. Other interesting directions include analyzing systems using other modulation schemes (instead of 2-code keying), or OCDMA spreading codes (instead of the standard Hadamard codes), or when *both* the scrambler setting and the codewords assigned to users are kept secret and refreshed periodically (*e.g.* codewords could be assigned from one of $W!$ instantiations of the Hadamard codes (see [2] §VII), where the instantiation chosen would be kept secret).

5 Conclusions

When analyzing the confidentiality provided by OCDMA systems, we have demonstrated the importance of formulating security analyzes using standard cryptanalytic notions (*e.g.*

Kerckhoffs' Principle, KPA-attacks), particularly if these systems are to present a viable alternative to standard digital encryption schemes. Moreover, the existence of the attacks we describe here suggest that *spectral-phase encoded OCDMA systems that use* only *the phase scrambler setting as a secret key are unlikely to guarantee confidentiality, unless the key is refreshed at rates comparable to the system data rates.* Other variants of OCDMA (*e.g.* when user codewords and scrambler settings are both secret) may or may not be secure. Determining the confidentiality of these variants is left for future work.

Acknowledgement

I'd like to thank Prof. Boaz Barak for his guidance and interest in discussing attacks on this system, Dr. Ron Menendez for helping me get my assumptions straight and for many helpful and motivating discussions, and Prof. Paul Prucnal for his support and interest in this project. I also appreciate the many technical discussions I've had with others here at Princeton, including Prof. Moses Charikar, Eugene Brevdo, Jimmy Chui and Shannon Hughes.

References

[1] P. R. Prucnal, Ed., *Optical Code Division Multiple Access: Fundamentals and Applications.* New York: Taylor and Francis, 2005.

[2] R. Menendez *et al.*, "Network applications of code translation for spectrally phase-encoded OCDMA," *J. Lightwave Technol.*, Oct. 2005.

[3] F. Xue, Y. Du, S. B. Yoo, and Z. Ding, "Security issues on spectral-phase-encoded optical CDMA with phase-masking scheme," *OSA Optical Fiber Communication Conference (OFC)*, 2006.

[4] T. Shake, "Security performance of optical CDMA against eavesdropping," *J. Lightwave Technol.*, vol. 23, no. 2, pp. 655–670, Feb. 2005.

[5] ——, "Confidentiality performance of spectral-phase-encoded optical CDMA," *J. Lightwave Technol.*, vol. 23, no. 4, pp. 1652–1666, Apr. 2005.

[6] N. Ferguson and B. Schneier, *Practical Cryptography.* Indianapolis, IN: Wiley, 2003.

[7] J. Jiang *et al.*, "Experimental investigation of security issues in OCDMA," *OFC 2006.*

[8] J. Nechvatal *et al.*, "Report on the development of the advanced encryption standard (aes)," *NIST*, Oct. 2000.

[9] A. Agarwal *et al.*, "Ring-resonator-based integrated photonic circuit for phase coherent applications," *J. Lightwave Technol.*, Jan. 2006.

A Proof of Claim 3.4

We restate Claim 3.4 for convenience.

Suppose the eavesdropper obtains any two arbitrary known plaintexts Θ_1, Θ_1 *and corresponding noise-free measurements* $\mathrm{y}_1, \mathrm{y}_2$. *Let* Θ_1, Θ_2 *be specified by diagonal* ± 1 *matrices* L_1, L_2 *as in Claim 3.2. Then if* any *of the following three matrices,* $L1$, L_2 *and* L_1L_2, *contain an odd number of '-1's then the eavesdropper immediately learns the key* k *by solving* $Q\widehat{\mathrm{k}}_a = 0$ *(i.e. since* $m = 1$*).*

We will prove this claim by showing that if *any* of the following three matrices: L_1, L_2 or L_1L_2 have an odd number of -1's along their diagonals then $rank(Q) = N - 1$. Recall that if $rank(Q) = N-1$ then $m = 1$ and there is only a single linearly independent vector $\widehat{\mathrm{k}}_a = \mathrm{k}_a$ satisfying $Q\widehat{\mathrm{k}}_a = 0$, so that the eavesdropper immediately learns the key.

We prove this claim in two steps. In Claims A.1-A.2 we show that $rank(Q) < N - 1$ (i.e. $m > 1$ and the eavesdropper "does not immediately learn the key") iff for $L = L_1$, $L = L_2$ and $L = L_1L_2$ the matrix $M = H_nLH_n$ satisfies an equation of the form

$$MZ_a = Z_bM \tag{14}$$

for Z_a, Z_b diagonal matrices with diagonal entries from $\{-1, 1\}$ where $Z_a, Z_b \neq \pm I$. Then in Claims A.4-A.5 we show that equation (14) cannot be satisfied for $Z_a, Z_b \neq \pm I$ if L contains an odd number of -1 entries along its diagonal. Combining these claims, it follows that a necessary (but not sufficient) condition for the adversary to *fail* to learn the key (i.e. $rank(Q) < N - 1$) is that all three of L_1, L_2 and L_1L_2 have an even number of -1 entries along their diagonals. Claim 3.4 follows from the contrapositive of this statement.

Claim A.1 *Given two plaintexts* $\Theta_1 = [H_N \; L_1H_N]$, $\Theta_2 = [H_N \; L_2H_N]$ *such that* $L_1 \neq L_2$ *and two corresponding measurements* $\mathrm{y}_1 = [\mathrm{y}_{a1} \; \mathrm{y}_{b1}]^T$, $\mathrm{y}_2 = [\mathrm{y}_{a2} \; \mathrm{y}_{b2}]^T$. *Then, form the matrix* Q *as in (13). If* Z *is a diagonal matrix with diagonal entries from* $\{-1, 1\}$, *then* $rank(Q) < N - 1$ *iff the matrices* $(H_NL_1L_2H_N)$, Z *commute, i.e.*

$$(H_NL_1L_2H_N)Z = Z(H_NL_1L_2H_N)$$

for some $Z \neq \pm I$.

Proof: First, we'll show why we are interested in commuting matrices. Recall that with two known plaintexts - measurement pairs, the eavesdropper has system of two matrix equations of the form in (10), namely

$$\begin{aligned} Y_{b1}^{-1} B_1^T A_1^{T^{-1}} Y_{a1} \widehat{\mathrm{k}}_a &= \widehat{\mathrm{k}}_b \\ Y_{b2}^{-1} B_2^T A_2^{T^{-1}} Y_{a2} \widehat{\mathrm{k}}_a &= \widehat{\mathrm{k}}_b \end{aligned} \tag{15}$$

Recall the $\widehat{\mathrm{k}}_a, \widehat{\mathrm{k}}_b$ are the eavesdroppers key guesses. Putting $A = H_N$ and $B = LH_N$ and rearranging the system of equation we obtain

$$\begin{aligned} H_N Y_{a1} \widehat{\mathrm{k}}_a &= L_1 H_N Y_{b1} \widehat{\mathrm{k}}_b \\ H_N Y_{a2} \widehat{\mathrm{k}}_a &= L_2 H_N Y_{b2} \widehat{\mathrm{k}}_b \end{aligned} \tag{16}$$

Now since the Y matrices are diagonal, we can swap the order of the Ys and the $\widehat{\mathrm{k}}$s to obtain

$$\begin{aligned} H_N \widehat{K}_a \mathrm{y}_{a1} &= L_1 H_N \widehat{K}_b \mathrm{y}_{b1} \\ H_N \widehat{K}_a \mathrm{y}_{a2} &= L_2 H_N \widehat{K}_b \mathrm{y}_{b2} \end{aligned} \tag{17}$$

where in our notation the $\widehat{K}$ matrices are also diagonal matrices. Now recall that $\mathrm{y}_a = K_a A^T \mathrm{x} = K_a H_N \mathrm{x}$ and $\mathrm{y}_b = K_b B^T \mathrm{x} = K_b H_N L \mathrm{x}$, where $K_a = diag(\mathrm{k}_a)$ and $\mathrm{k}_a, \mathrm{k}_b$ is the true secret key used by the system (and unknown to the eavesdropper). Substitution for $\mathrm{y}_{a1}, \mathrm{y}_{b1}, \mathrm{y}_{a2}, \mathrm{y}_{b2}$ into (17) gives

$$\begin{aligned} H_N \widehat{K}_a K_a H_N \mathrm{x}_1 &= L_1 H_N \widehat{K}_b K_b H_N L_1 \mathrm{x}_1 \\ H_N \widehat{K}_a K_a H_N \mathrm{x}_2 &= L_2 H_N \widehat{K}_b K_b H_N L_2 \mathrm{x}_2 \end{aligned} \tag{18}$$

which implies that

$$\begin{aligned} H_N \widehat{K}_a K_a H_N &= L_1 H_N \widehat{K}_b K_b H_N L_1 \\ H_N \widehat{K}_a K_a H_N &= L_2 H_N \widehat{K}_b K_b H_N L_2 \end{aligned} \tag{19}$$

now equating the right sides of the equations above, and premultiplying by $H_n L_1$ and postmultiplying by $L_1 H_n$ gives

$$H_N L_1 L_2 H_N \widehat{K}_b K_b = \widehat{K}_b K_b H_N L_1 L_2 H_N$$

and if we let $Z = \widehat{K}_b K_b$ we have that

$$(H_N L_1 L_2 H_N) Z = Z (H_N L_1 L_2 H_N) \tag{20}$$

We switch gears to understand to implications of (20). Consider an eavesdropper that obtains two plaintexts such that $L_1 \neq L_2$, and two corresponding measurements $\mathrm{y}_{a1}, \mathrm{y}_{b1}, \mathrm{y}_{a2}, \mathrm{y}_{b2}$.

First, we observe that since Q was formed using two measurements $\mathrm{y}_1, \mathrm{y}_2$ that correspond correctly to two plaintexts L_1, L_2, it follows that the system of equations (16) must have a solution - namely that $[\widehat{\mathrm{k}}_a, \widehat{\mathrm{k}}_b] = [\mathrm{k}_a, \mathrm{k}_b]$. Equivalently, the equation $Q\widehat{\mathrm{k}}_a = 0$ must have a solution $\widehat{\mathrm{k}}_a = \mathrm{k}_a$. It follows that $rank(Q) < N$, since if $Q\widehat{\mathrm{k}}_a = 0$ has a solution in $\{1, -1\}^N$ then the nullspace of Q has dimension at least unity.

Now when $rank(Q) = N - 1$ then the only solutions to $Q\widehat{\mathrm{k}}_a = 0$ for $\widehat{\mathrm{k}}_a \in \{-1, 1\}^n$ are $\widehat{\mathrm{k}}_a = \pm\mathrm{k}_a$. Then it follows that when $rank(Q) = n - 1$, then $\widehat{K}_a = \pm K_a$ so that $Z = \pm I$.

Now we are particularly interested in cases when the adversary learns many possibilities for the key k_a (i.e. "does not immediately learn the key") by solving $Q\widehat{\mathrm{k}}_a = 0$. Equivalently, we are interested in cases where $rank(Q) < N - 1$. In such cases, it follows that there exists $\widehat{\mathrm{k}}_a \in \{-1, 1\}^n$ such that $\widehat{\mathrm{k}}_a \neq \pm\mathrm{k}_a$. Equivalently, there must exist $Z = \widehat{K}_a K_a \neq \pm I$ satisfying (20), which is exactly the statement of this claim.

Finally, we mention why we restricted the claim to cases when $L_1 \neq L_2$. Suppose instead that $L_1 = L_2$. Then, $H_N L_1 L_2 H_N = H_N I H_N = NI$ (since H_N is a symmetric Hadamard matrix) that commutes with any arbitrary diagonal matrix with diagonals from $\{-1, 1\}$, even if $rank(Q) \geq N - 1$, contradicting our claim. ∎

Claim A.2 *Given two known plaintexts (L_1, L_2) and two corresponding measurements $\mathrm{y}_1 = [\mathrm{y}_{a1}\ \mathrm{y}_{b1}]^T$, $\mathrm{y}_2 = [\mathrm{y}_{a2}\ \mathrm{y}_{b2}]^T$. Then, form the matrix Q as in (13). If Z_a, Z_b are two diagonal matrices with diagonal entries from $\{-1, 1\}$, then $rank(Q) < N - 1$ iff for $L = L_1$ and $L = L_2$ then*

$$(H_N L H_N) Z_a = Z_b (H_N L H_N) \tag{21}$$

such that $Z_a \neq \pm I$.

Proof: As in Claim A.1, we rewrite the system of equations in (15) as (19), and rearranging we obtain

$$\begin{aligned} H_N L_1 H_N \widehat{K}_a K_a &= \widehat{K}_b K_b H_N L_1 H_N \\ H_N L_2 H_N \widehat{K}_a K_a &= \widehat{K}_b K_b H_N L_2 H_N \end{aligned} \tag{22}$$

Therefore, letting $Z_a = \widehat{K}_a K_a$ and $Z_b = \widehat{K}_b K_b$ we arrive at the condition in (21)

We a similar argument to the one in the proof of Claim A.1 to explain the significance of (21). First, if Q was formed using two measurements $\mathrm{y}_1, \mathrm{y}_2$ that correspond correctly

to two plaintexts L_1, L_2, it follows that $rank(Q) < N$, since $Q\widehat{\mathrm{k}}_a = 0$ has solution $\widehat{\mathrm{k}}_a = \mathrm{k}_a$. Now when $rank(Q) = N - 1$ then the only solutions to $Q\widehat{\mathrm{k}}_a = 0$ for $\widehat{\mathrm{k}}_a \in \{-1, 1\}^N$ are $\widehat{\mathrm{k}}_a = \pm\mathrm{k}_a$. Equivalently, when $rank(Q) = N - 1$, then $\widehat{K}_a = \pm K_a$ so that $Z_a = \pm I$. Furthermore, substitution of $\widehat{K}_a = \pm K_a$ into (22) gives $\widehat{K}_b = \pm K_b$ so that $Z_b = \pm I$. Now when $rank(Q) < N - 1$ it follows that there exists $\widehat{\mathrm{k}}_a \in \{-1, 1\}^N$ such $Q\widehat{\mathrm{k}}_a = 0$ and $\widehat{\mathrm{k}}_a \neq \pm\mathrm{k}_a$. Equivalently, $\widehat{K}_a \neq \pm K_a$ so that there must be $Z_b \neq \pm I$ satisfying (21) which is exactly the statement of this claim. ∎

Before we go on, we note that Claims A.1-A.2 give rise to following Lemma, since they show the equivalence between the condition $rank(Q) < N - 1$ and a condition that depends only on (L_1, L_2). Equivalently the Lemma shows that $rank(Q)$ and the eavesdroppers search space is completely independent of the choice of key $\mathrm{k}_a, \mathrm{k}_b$ and the inter-user phases $\mathrm{x}_1, \mathrm{x}_2$.

Lemma A.3 *Given two known plaintexts (L_1, L_2) and two corresponding measurements $\mathrm{y}_1 = [\mathrm{y}_{a1}\ \mathrm{y}_{b1}]^T$, $\mathrm{y}_2 = [\mathrm{y}_{a2}\ \mathrm{y}_{b2}]^T$. Then, form the matrix Q as in (13). Then whether or not $rank(Q) < N - 1$ is completely determined by (L_1, L_2).*

We now return to our main proof:

Claim A.4 *Let M be any matrix and Z_a, Z_b be diagonal matrices with diagonal entries on $\{-1, 1\}$. Then if*

$$MZ_a = Z_b M \tag{23}$$

then $M_{ij} = 0$ for i, j such that $a_{ii} = -b_{jj}$.

Proof: Writing $MZ_a = Z_b M$ in matrix form as

$$\begin{bmatrix} M_{11} & \dots & M_{1n} \\ \vdots & & \vdots \\ M_{n1} & \dots & M_{nn} \end{bmatrix} \begin{bmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{bmatrix} = \begin{bmatrix} b_{11} & & \\ & \ddots & \\ & & b_{nn} \end{bmatrix} \begin{bmatrix} M_{11} & \dots & M_{1n} \\ \vdots & & \vdots \\ M_{n1} & \dots & M_{nn} \end{bmatrix} \tag{24}$$

and comparing element by element, we can see that for all i, j we must have

$$M_{ij} b_{ii} = M_{ij} a_{jj}$$

and when $a_{ii} \neq b_{jj}$ (i.e. $a_{ii} = -b_{jj}$) we must have that $M_{ij} = 0$. ∎

Claim A.5 *Let L, Z_a, Z_b be a diagonal matrices with diagonal entries from $\{-1, 1\}$. Then if L has an odd number of -1 entries along its diagonal, the the only solution to*

$$(H_N L H_N) Z_a = Z_b (H_N L H_N) \tag{25}$$

is $Z_a = Z_b = \pm I$.

Proof: First, since L is a diagonal matrix with diagonal entries from $\{-1, 1\}$, we can rewrite L as

$$L = I - 2 \sum_{\ell_i} diag(e_{\ell_i})$$

where ℓ_i are the location of the -1 entries along the diagonal of L, and e_ℓ is a the ℓ^{th} standard basis vector. Then, it follows that

$$\begin{aligned} M &= H_N L H_N \\ &= H_N (I - 2 \textstyle\sum_{\ell_i} diag(e_{\ell_i})) H_N \\ &= H_N H_N - \frac{2}{n} \textstyle\sum_{\ell_i} H_N diag(e_{\ell_i}) H_N \\ &= NI - 2 \textstyle\sum_{\ell_i} h_{\ell_i} h_{\ell_i}^T \end{aligned} \tag{26}$$

where h_ℓ is the ℓ^{th} column of the Hadamard matrix H_N. Since h_ℓ is a $\{+1, -1\}$ vector for all ℓ, it is easy to see that the dyad matrix $h_{\ell_i}^T h_{\ell_i}$ will also be an $N \times N$ matrix of $\pm$1s.

Now from Claim A.4, we know that if M commutes with a diagonal matrices Z_a, Z_b with diagonal entries from $\{-1, 1\}$ where $Z_a, Z_b \neq \pm I$, it follows that M must have at least two off-diagonal entries that are equal to zero. However from (26), we have that any off-diagonal entry of M has the form

$$M_{ij} = 0 - \sum_{\ell_i} \alpha_{ij} \tag{27}$$

where $\alpha_{ij} = \pm 1$. Now where there are an odd number of negative -1's in the $L_1 L_2$ matrix, it follows that there are an odd number of terms in the sum in (27) , and it follows that $M_{ij} \neq 0$ for all (i, j). Thus we have arrived at a contradiction, and M cannot satisfy (25) with diagonal matrices with entries from $\{-1, 1\}$ other than $Z_a = Z_b = \pm I$. ∎