# The Need for Media Access Control in Optical CDMA Networks

Purushotham Kamath, Joseph D. Touch and Joseph A. Bannister
USC/ISI
4676 Admiralty Way
Marina Del Rey, CA 90292
{pkamath, touch, joseph}@isi.edu

***Abstract*— Optical CDMA Local Area Networks allow shared access to a broadcast medium. Every node on the network is assigned an Optical Orthogonal Codeword (OOC) to transmit or receive on. OOCs are designed to be pseudo-orthogonal, i.e., the correlation (and therefore the interference) between pairs of codewords is constrained. This paper demonstrates that the use of optical CDMA does not preclude the need for a media access control (MAC) layer protocol to resolve contention for the shared media.**

**OOCs have low spectral efficiency. As more codewords are transmitted simultaneously, the interference between codewords increases and the network throughput falls. This paper analyzes a network architecture where there is virtually no MAC layer, except for choice of the codeset, and shows that its throughput degrades and collapses under moderate to heavy load. We propose an alternate architecture called *Interference Avoidance* where nodes on the network use media access mechanisms to avoid causing interference on the line, thereby improving network throughput. Interference avoidance is analyzed and it is shown that it can provide up to 30% improvement in throughput with low delays and no throughput collapse. We validate our analysis through simulation with realistic network traffic traces.**

**KEYWORDS: System design**

## I. INTRODUCTION

Code Division Multiple Access (CDMA) has been widely used in wireless networks such as the cellular phone system for several years because of its resilience to multiuser interference. Its use on an optical link has been studied extensively [1], [2], [3]. However, several concerns have been expressed about the use of spread spectrum on an optical link. The main concern is that the use of optical CDMA results in low network throughput [4].

Code division multiplexing on an optical link is significantly different from that on the wireless medium. The primary difference is that optical fiber, in contrast to wireless, is an intensity medium (also called a unipolar or incoherent medium). Signals are transmitted as optical power. Hence, binary data is sent using pulses of light. The overlap of optical pulses results in the addition of optical power.

[1]This material is based upon work supported by the Defense Advanced Research Projects Agency under contract no. N66001-02-1-8939 issued by the Space and Naval Warfare Systems Center (SPAWAR). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Defense Advanced Research Projects Agency, SPAWAR, or the U.S. Government.

An Optical Orthogonal Code (OOC) set is a set of (0,1) sequences of length $N$ that satisfies certain autocorrelation and crosscorrelation constraints. The term *codeset* is used to refer to the set of such sequences, and the term *codeword* is used for a member of the set. Each 0 or 1 of a sequence is called a *chip*, and the sequence represents a data *bit*. The number $w$ of 1 chips of a codeword of the codeset is called its Hamming weight. This paper considers *constant weight* codesets, i.e., codesets with all codewords having the same weight.

In ON-OFF keyed (OOK) optical CDMA networks, data is sent using codewords. To send a 1 data bit, the node transmits the codeword and the send a 0 data bit, it does not transmit any code. The presence of the codeword signifies a data bit of 1 and the absence signifies a 0 data bit.

Let $s(t)$ denote the value of the $t^{th}$ chip of an $N$ chip codeword $s$. For any codeword $s_1$ in the codeset, the autocorrelation constraint is

$$\sum_{t=0}^{N-1} s_1(t+\tau)s_1(t) \quad \begin{array}{ll} = w & \textit{when} \ \ \tau = 0 \\ \leq \kappa_a & \textit{when} \ \ 0 < \tau \leq N-1 \end{array}$$

For any pair of codewords $s_1$ and $s_2$ in the codeset, the crosscorrelation constraint is

$$\sum_{t=0}^{N-1} s_1(t+\tau)s_2(t) \quad \leq \kappa_c \qquad \textit{when} \ \ 0 \leq \tau \leq N-1$$

A codeset designed under these constraints is said to be *pseudo-orthogonal*. The pseudo-orthogonality constraint ensures that the codesets have limited interference between codewords. Most codeset designs use $\kappa_a = \kappa_c$ and call this value the *Maximum Collision Parameter* $\kappa$. Most optical CDMA networks are designed to use $\kappa = 1$ or 2, to ensure that interference between codewords is low. A particular codeset is specified by the parameters $(N, w, \kappa)$.

The *size* of a codeset is the number of codewords in the codeset. The size $S$ of a constant weight code constructed under the pseudo-orthogonality constraints can be shown to satisfy the Johnson bound [5], [6]:

$$S(N, w, \kappa) \leq \left\lfloor \frac{1}{w} \left\lfloor \frac{n-1}{w-1} \left\lfloor \frac{n-2}{w-2} \left\lfloor ..... \left\lfloor \frac{n-\kappa}{w-\kappa} \right\rfloor \right\rfloor \right\rfloor .... \right\rfloor \right\rfloor$$

Codesets are designed to be pseudo-orthogonal for low interference. Therefore multiple nodes can transmit simultaneously on different codewords. A receiver that is tuned to a particular codeword listens in the appropriate chip locations where it expects to receive ‘1 chips’. If the number of ‘1 chips’ its detects is above a certain *threshold*, it assumes that it has received a ‘1’ data bit. The threshold is generally set to the weight of the code. An error will occur if there are enough other codewords on the line which have ‘1 chips’ in the same locations as the expected codeword. If that happens, a ‘0’ data bit or no transmission may be falsely detected as a ‘1’ data bit (*a false positive*). If this happens it will result in the loss of the data packet. This is the main cause for the degradation in throughput of an optical CDMA network at high offered load.

The *spectral efficiency* of a codeset is given by

$$\eta = A/N$$

where $A$ is the number of active (or simultaneous) users. The spectral efficiency is a measure of how good the coding scheme performs compared to a perfect time division multiplexed system, for which $\eta = 1$.

The use of optical CDMA with a threshold based detector puts a limit on the amount of bandwidth available. For example, the Johnson bound for a (25, 3, 1) codeset gives a maximum of $\lfloor 1/3 * \lfloor (25-1)/(3-1) \rfloor \rfloor = 4$ codewords. At a chipping rate of 1Gb/s, this means that the data rate is (1/25)Gb/s = 40Mb/s. Under worst case conditions, for zero false positive errors, a maximum of 3/1 = 3 codewords can be on the line simultaneously. The spectral efficiency is just $3/25 = 12\%$. Thus, only a fraction of the available bandwidth is utilized. In the above example, even under the best case conditions the available bandwidth is approximately $1/w = 1/3$ of 1Gb/s $\approx$ 300Mb/s. This is because we are using 3 chips to represent what could have been sent using 1 chip. Attempting to increase the bandwidth utilization by increasing the number of codewords on the line increases the interference between codewords resulting in lost packets and lower network throughput.

To summarize, optical CDMA allows nodes to transmit asynchronously without any media access delay. However, it has two disadvantages: low spectral efficiency and low throughput under heavy loads. At high offered loads, the cause for low throughput is the interference between codewords.

In Section II we further discuss the motivation for our study. Related work in this field is discussed in Section III. Section IV outlines the system design. The media access mechanisms are described in detail in Section V. Analysis of an optical CDMA system with and without interference avoidance follows in Sections VI and VII. The limitations of the architecture and conclusions are discussed in Sections IX and X.

## II. Motivation

The throughput of an optical CDMA network at any instant of time depends on the codewords that are on the line at that instant. The *chip offset* between any two codewords is defined as the difference in chip times between the start of transmission of the codewords. The interference depends on the exact codewords on the line and the chip offsets with respect to each other. For example, consider the codewords shown in Figure 1(a). The analysis assumes that codewords are chip aligned but not necessarily bit aligned. The figure is a snapshot of data bits sent by two nodes (codewords 1 and 2). A third node is preparing to transmit on codeword 3. The figure represents an instant in time when all three packets had a 1 data bit on the line. Only one data bit from each node is shown in the figure. The packets could contain several 0 and 1 data bits, extending in both directions. The packets sent on codewords 1 and 2 can be transmitted without any problem under the chip offsets shown. However if a packet with codeword 3 (the codeword to be transmitted) were to be transmitted with the chip offset shown, there is a high probability that it would not be received properly. This is because it is likely that the packet sent on codeword 3 would contain a 0 data bit. Codewords 1 and 2 could potentially combine to create a 1 data bit in that position. The checksum on the packet would fail and it would be discarded.

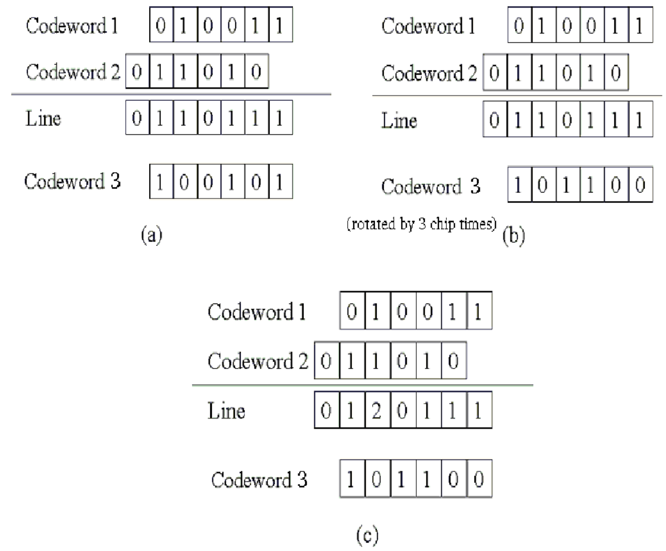


Fig. 1. Examples of interference between codewords

On the other hand, if the packet on codeword 3 was sent three chip times later as shown in Figure 1(b), the three packets could be transmitted without interfering with each other. The figure shows the codeword 3 shifted by 3 chip times (it shows the end of one data bit and the start of another). Codeword 3 has at least one chip that does not interfere with codewords 1 and 2. Hence it will be received correctly.

The main contribution of this work is to recognize that the throughput of optical CDMA under heavy loads can be improved by simple media access mechanisms that prevent interfering codewords from being sent simultaneously. The media access mechanism does this by transmitting the packet at the appropriate chip offset with respect to the codewords on the line. It senses the interference on the line and decides when to transmit. We call this scheme *Interference Avoidance*.

The sensing mechanism is feasible even though the chipping rate of the network is high (say, in the order of Gb/s) because the data rate is lower (in the order of Mb/s). This allows the use of well known media access techniques such as carrier sensing and collision detection. It may be noted that though the data rate is low, the overall network throughput is high (in the order of Gb/s), because several nodes transmit simultaneously. The Interference Avoidance mechanism is analogous to the well known Carrier Sense Multiple Access/Collision Detection (CSMA/CD) mechanism in several ways, though it has several key differences.

An alternative method to improve the throughput of an optical CDMA network is to vary the codeset design parameters (the codeword length $N$ or the maximum crosscorrelation parameter $\kappa$). Increasing $\kappa$ allows the codeset to have more codewords. The cross correlation between the codewords varies between 0 and $\kappa$. Given the random nature of transmitted codewords, it is possible that the pairs of codewords with maximum cross correlation appear on the line only rarely and this may improve throughput. Through analysis, we show that this method can provide only limited gains in overall network throughput and that for higher gains a media access protocol is necessary.

It may be noted that the problem we attempt to solve is significantly different from the problem on wireless media where the nature of the codes is different, where hidden terminals and exposed terminals prevent effective sensing operations and multipath effects degrade throughput.

To understand the performance implications of Interference Avoidance our analysis tries to answer the following questions:

- Can throughput improvements be obtained by using codeset design techniques (varying $N$, $\kappa$) without using interference avoidance?
- How much improvement in throughput does interference avoidance provide and what is the cost in terms of increased latency?
- What are the tradeoffs that this design provides?

We analyze interference avoidance and show that it can produce significant throughput gains (up to 30%) with low delay under certain conditions.

## III. Related work

This work draws on several earlier results, both from the area of code design and the area of network design.

Chung et al. [7] describe several algorithms to construct OOCs. These constructions are for codes with maximum crosscorrelation parameter $\kappa = 1$. Chung and Kumar [5] describe a method for construction of codes with $\kappa = 2$. Several construction methods for OOCs are described in [8] and [9] among others. The construction methods focus on codes with low crosscorrelation parameter.

In the area of optical CDMA network design, Salehi [2], [3] analyzed an optical CDMA based network and developed expressions for the bit error rate of a network that uses codesets with $\kappa = 1$. Azizoglu et al. [10] determined the error rate for codesets with $\kappa = 2$. They showed that the bit error rate does not degrade significantly when $\kappa$ is increased from 1 to 2.

Recently, Shalaby [11], [12] examined the effect of two different crosscorrelation parameters on the throughput of an OOK-CDMA network. The result of the analysis was that under certain conditions, throughput could be increased by a factor of around 3 by increasing the $\kappa$ from 1 to 2.

Hsu et al. [13], [14] analyzed the performance of slotted and unslotted optical CDMA packet networks. They developed expressions for the throughput of the network and showed performance can be improved using Forward Error Correction (FEC) codes and hard limiters. Lee et al. [15] analyzed the performance of OOCs by assuming crosscorrelation distributions to be Gaussian. They used their analysis to compare the performance of different code constructions. They showed that the performance of optical CDMA networks depends on the mean and variance of the crosscorrelation values.

Muckenheim et al. [16], [17] studied the effect of bit error probability on the packet error probability and suggested the use of block codes to reduce packet errors. They also described a random delay protocol to reduce the errors incurred during periods of high activity.

Kumar [18] analyzed the stability and throughput of optical CDMA networks using various protocols. They showed how the saturation throughput degrades with code sharing. In the context of packet radio networks, Raychaudhuri [19] analyzed the throughput of a generic CDMA based packet switched network. Polydoros [20] analyzed the performance of a random access spread spectrum network.

MAC mechanisms such as Carrier Sense and Collision Detection have been explored in the context of optical networks in [21]. The technical details of implementing a carrier sense/collision detection mechanism over an optical medium have been discussed.

The approaches discussed above attempt to improve throughput through either code design, the use of optical devices or through system design. In contrast, the approach we propose, interference avoidance, (discussed in Section V) is a media access control mechanism.

## IV. System architecture

The following sections describe the network architecture, the addressing, code allocation mechanism and the design of a receiver. The interference avoidance technique is independent of some architecture details such as the addressing and the codeword allocation scheme. Other architectural decisions such as the use of a star coupler and the receiver design are related to the use of the interference avoidance mechanism. The complete architecture is included here to present a clear picture of how the network will function.

### A. Network architecture

The network we describe is a broadcast star coupler based system. Star couplers are passive optical elements with all inputs connected to all outputs. Data transmitted on an input of the coupler is transmitted to all its outputs. A port on the coupler consists of an input and an output. Star couplers typically have between 2 to 128 ports. Every node on the network is connected to the coupler. Every node is equipped with at least one transmitter and one receiver. The transmitter and receiver may be tuned to any codeword. Though the architecture assumes that there is only one wavelength available, it can be used with Wavelength Division Multiplexing (WDM). When more than one wavelength is available, the protocol may be run separately on each wavelength. For reasons of simplicity we consider an architecture with one wavelength.

### B. Addressing and packet format

Every node has a unique *node address* which is distinct from the codewords in the codeset. This address is permanent and is unique across the network. The packet header has a preamble to allow nodes to detect the start of a packet and an error detection mechanism such as a checksum to detect corrupted packets. The packet header also has a length field that specifies the length of the packet. An encoding of the data packet such as 4B/5B is used to ensure that long sequences of either 0s or 1s are prevented.

### C. Codeword allocation

A node can choose to transmit or receive on any codeword. For simplicity, we assume a tunable transmitter-fixed receiver system. Nodes can tune their transmitters to transmit on any code and their receivers are fixed to receive on a single code. Nodes choose which codeword to receive on when they start up. The codeword chosen is a hash of the node address. The hash function is known to all nodes, so a node that wishes to transmit to another node can determine the codeword on which to transmit using the node address of the receiver and the hash function. This removes the need for a control channel or a centralized server to perform the mapping of node address to codeword. Although this is not a requirement of our design, for reasons of simplicity we assume that this is how the system is designed.

In contrast, several optical CDMA network architectures use a static allocation of codewords to nodes. Each node on the network is assigned a unique codeword. This method restricts the design of the codeset because it should be large enough to support the number of nodes.

Our design means that nodes will share codewords, i.e., several nodes may receive on the same codeword at a time. Any codeword or any of its cyclic shifts may be on the line at any time. Nodes accept or discard packets they receive based on the node address in the packet header.

### D. Receiver design

As mentioned in Section I, false positives are the source of errors: a 1 bit is detected when a 0 bit is being sent. There are two error cases to be considered when designing a receiver:

- False positives detected on a codeword when that codeword is not being transmitted may be detected by the absence of a preamble or by a checksum failure in the packet.
- False positives detected on a codeword when that codeword is being transmitted will result in the 0 data bits of the packet being detected as 1s. (The 1 data bits will not be corrupted.) If this happens, it is possible that the packet might be corrupted. A packet on an average will consist of an equal number of 0 and 1 bits. If a 4B/5B encoding is used, a 0 data bit will occur at least every 4 bits. Therefore if this case occurs, we assume that this packet is lost with probability 1.

Thus, if any combination of codewords on the line add up to another packet's codeword, then that packet is lost. The event is called a *bit collision*.

Each receiver is tuned to a particular optical CDMA codeword. It continuously listens for that codeword and as soon as it successfully detects a data bit and the packet preamble, it continues to listen for a packet and performs a checksum operation on the packet once it has been completely received. If two nodes transmit to a single node at around the same time, the receiver receives the first packet and synchronizes to it.

We assume that the receivers do not do any form of power limiting. The network is a broadcast network and all nodes see all transmissions. We assume that every node sees exactly the same data on the line, possibly after different propagation delays. We also assume that the effect of different fiber lengths may be taken care of by using calibration mechanisms that allow each node to measure its propagation delay from the node to the coupler with the granularity of a chip time.

## V. Media access

We describe two forms of media access, *Aloha-CDMA* and *Interference Sense/Interference Detection (Is/Id)*, our proposed mechanism.

### A. Aloha-CDMA

This is the conventional form of access in optical CDMA networks. There is no explicit media access protocol. Nodes can transmit asynchronously with no media access protocol. This is similar to unslotted Aloha. The codeset used may be chosen to maximize throughput. The parameters used to construct the codeset used may be varied to control the interference between codewords. In Section VI we show analytically that the throughput degrades as offered load increases and is low irrespective of the parameters of the codeset used.

### B. Interference sense and Interference detection (Is/Id)

The main reason that the throughput of the network degrades with the Aloha-CDMA mechanism is that packets are sent without sensing the media. Codewords with high interference between them may be sent on the line at the same time.

To improve the throughput, we use mechanisms similar to the well known media access mechanisms of carrier sense and collision detection. Carrier sense and collision detection and their various flavors (non-persistent, $p$-persistent, etc.) have been analyzed in [22] and elsewhere.

In *Is/Id* a node senses if its transmission would cause interference before transmission. If it senses that interference would occur, it defers and attempts to transmit again after a delay. This is called *Interference Sense*. This form of media access means that a node must have at least two receivers, if it wants to transmit and receive at the same time.

After sensing the line, a node must decide whether to transmit or not. There are two cases to consider:

- Will the current state of the line impact the node's transmission?
- Will the node's transmission impact other codewords on the line?

It is possible for the node to sense whether the transmission on the line will interfere with its transmission by comparing the chips on the line with its codeword. For example, in Figure 1(a) the transmission on the line (codeword 3) has power in the 1st, 4th and 6th chips and will therefore overlap with the '1 chips' of the codeword to be transmitted. If the transmission were delayed by 3 chip times, then there would be no interference.

The node needs a non-limiting receiver to determine if its transmission will impact other codewords on the line. A non-limiting receiver is able to determine the magnitude of each overlap by sensing the total power level. For example, in Figure 1(c), the node that is to transmit knows that there is potential for a codeword to be lost, because it sees that the addition of its code will result in 3 overlaps. It knows that the weight of the codeword is 3 and so it knows that there is potential for interference. This does not necessarily mean that there will be interference, but the probability of interference increases with the number of such overlaps. With a hard limiting receiver it is not possible to determine this as indicated in Figure 1(b).

To limit the interference on the line, we define two *interference sensing parameters*:

- The maximum magnitude of an overlap, the *overlap magnitude limit* $thresh_m$
- The number of the maximum magnitude overlaps, the *overlap count limit* $thresh_c$

Before transmission, a node determines whether its transmission would cause these limits to be exceeded. If it does, the node does not transmit and backs off. Alternatively, instead of backing off, the node could choose to transmit a few chip times later, if it determines that delaying the packet by a few chip times will reduce chances of interference.

After starting transmission, a node continues to sense for interference. Due to the finite propagation delay, the packet still remains vulnerable to transmissions from other nodes that may have been started in the interval between the start of transmission and the packet reaching them. The other nodes may not yet have sensed the sender's transmission. We term the interval during which this could happen, the *vulnerable period*. During transmission, if the sending node determines that interference has occurred, it can choose to stop transmission, back off and retransmit. This is called *Interference Detection*.

Several mechanisms have been studied in the context of carrier sense and collision detection in CSMA/CD networks to reduce the delay and to avoid capture of the medium [23]. In particular, there are two parameters which influence the delay: the interval after which the node retries, called the *backoff time*, and the number of times that the node attempts to retransmit a packet, called the *backoff count*.

Reducing the interference sensing parameters will limit the number of codes allowed on the line and thereby reduce interference. The parameters can be tuned for minimum interference or to allow a certain amount of interference. An ideal mechanism would adjust the limits such that the number of codes on the line is maintained optimal and the media access delays are kept within bounds.

The interference sensing operation takes a finite amount of time. With 4B/5B encoding of data, five data bit times will be sufficient to receive information about all the codes currently being transmitted on the line. For a network of chipping rate 1Gb/s and $N = 100$, the data rate is 10Mb/s. This implies that the sensing operation must be done within a few tenths of a microsecond which is well within the capability of today's processors.

It is interesting to note that the media access delay does not necessarily mean that packets suffer queuing delays. Packets can be transmitted out of order. For example, a node may have to transmit two packets on two different codewords. Interference on the line may prohibit the sending of the first packet, but may allow the second packet to be transmitted.

A more formal discussion of different algorithms for interference sensing and detection may be found in [24]. A deeper analysis of the sensing parameters and their effect on throughput may also be found there.

## VI. Analysis of Aloha-CDMA with different codeset parameters

In this section we analyze the throughput of the Aloha-CDMA mechanism under different code construction parameters. We show that varying the codeset design parameters (increasing the code length or relaxing the correlation constraints) does not provide much benefit in terms of network throughput.

The approach we follow is to develop expressions for the bit collision probabilities under different code design parameters and then extend these expressions to determine the network throughput.

The correlation constraint $\kappa$ influences the amount of interference on the network. A codeset may be constructed to minimize the interference between codeword pairs by minimizing $\kappa$. This results in a low number of codewords in the codeset. An alternative design is to increase the codewords in a codeset by reducing the constraints on $\kappa$. Depending on the distribution of the correlation between codewords, it is possible that this may result in a reduction in interference and an increase in throughput.

The values $\kappa = 1$ and $\kappa = w$ represent the bounds within which a codeset may be constructed. They provide lower and upper bounds on the maximum crosscorrelation of an OOC. A codeset with $\kappa = w$, will have several codes whose chip offsets correspond to other codes. This codeset represents the upper bound on the maximum crosscorrelation between individual codewords. To determine the impact of varying $\kappa$ on the bit collision probability we develop expressions for the bit collision probability when $\kappa = 1$ and $\kappa = w$.

### *A. Probability of bit collision for $\kappa = w$*

When $\kappa = w$ there is no constraint on the placement of 1's in a codeword, except for the constant weight constraint. 1's may be placed anywhere in the codeword because the restriction is that maximum number of common 1's between any pair of codes is equal to the weight of the code.

An error occurs when the number of overlaps is equal to the receiver threshold. The threshold is generally set to the weight $w$ of the code. For a constant weight codeset of length $N$, weight $w$, if any two codewords are chosen randomly from the sample space consisting of the codewords and their cyclic shifts, then the probability of them having $b$ chips overlapping is given by the *hypergeometric* distribution [25]. The probability of a bit collision, given that there are $m$ packets on the line, in addition to the codeword being received, $P_m(w)$, is calculated in Appendix I.

### *B. Probability of bit collision for $\kappa = 1$*

The maximum cross correlation between pairs of codes is 1. Therefore, if any two codewords are chosen randomly from the sample space consisting of the codewords and their cyclic shifts, then there are three cases: either 1 chip may overlap, 0 chips may overlap or $w$ chips may overlap (if the same codeword is chosen). The bit collision probability given $m$ packets on the line in addition to the codeword being received, $P_m(w)$ is calculated in Appendix II.

### *C. Bit collision probability vs. offered load*

If there are $m$ bits simultaneously on the line, then the offered load on the network is $mB/N$, where $B$ is the chipping rate of the network. Expressed as a fraction of the chipping rate, the normalized offered load is $m/N$.

A graph of the probability of bit collision against the offered load is shown in Figure 2. The graph shows the bit collision probability for four typical code sets. The collision probability is marginally lower for a ($N$, $w$, $\kappa$) = (10, 3, 3) codeset compared to a (10, 3, 1) codeset indicating that our hypothesis that an increase in $\kappa$ may result in overall lower bit collision probability may be correct. However this improvement is marginal and reduces as the load on the network increases. However for longer codes, such as the (100, 3, 1) and (100, 3, 3) where $N \gg w$, the hypothesis is incorrect. There is no discernible difference in the bit collision probabilities. The bit collision probability for a 100 chip codeset is lower than that of a 10 chip codeset. However this does not necessarily translate into higher network throughput as will be shown.

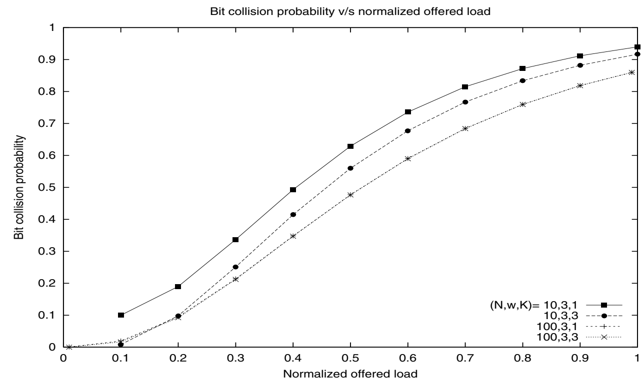


Fig. 2. Probability of a bit collision ($P_m(w)$) against normalized offered load ($m/N$) for different codeset design parameters ($N$, $w$, $\kappa$) for Aloha-CDMA.

### *D. Network throughput*

The presence of simultaneous packets on the line in Aloha-CDMA can be modeled by an $M/M/\infty$ queue. We choose the Poisson traffic model for ease of calculation. We consider more realistic traffic models in later sections and validate our hypothesis through simulation.

Consider an infinite user population and let the aggregate traffic arrival rate be Poisson with an average packet arrival rate of $\lambda$ packets per second. Assume the packet lengths are exponentially distributed with an average packet length of $L$ bits. When a packet is transmitted it stays on the line for a duration of time that is exponentially distributed with average equal to the packet length divided by the data rate in bits per second. Thus,

$$1/\mu = L/(B/N)$$

where $\mu$ is the average service rate, $B$ is the overall network bandwidth (the chipping rate) and $N$ is the code length.

Packets are transmitted on the line on arrival and there is no limit on the number of simultaneous packets on the line. Therefore, the number of packets on the line is equivalent to the number of packets in an $M/M/\infty$ system. From queuing analysis [26], the probability of having $n$ packets in an $M/M/\infty$ system (on the line) is

$$P_{line}(n) = \frac{\rho^n e^{-\rho}}{n!}$$

where $\rho = \lambda/\mu$. An error occurs when there are $n$ packets on the line that can interfere with the code being received. The probability of a packet collision for an arrival rate $\lambda$, packet length $L$, codeword length $N$ is given by

$$P_{error} = \sum_{n=0}^{\infty} P_{line}(n).P_n(w)$$

where $P_n(w)$ is the probability of an error given that $n$ bits (codewords) are simultaneously on the line (from Section VI-A). The throughput efficiency of the network, the fraction of packets that are received correctly, is given by

$$Th_{aloha-cdma} = 1 - P_{error}$$

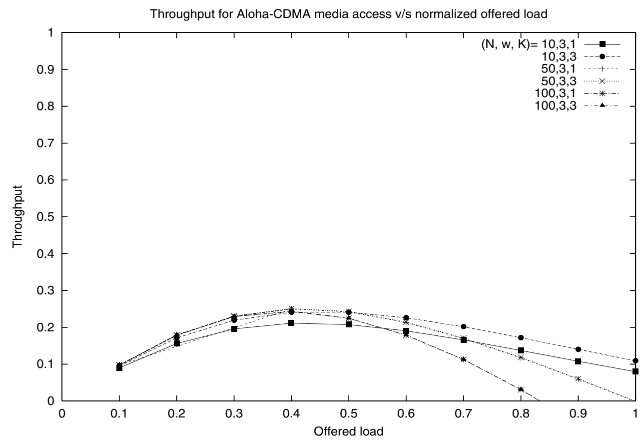


Fig. 3. Network throughput ($Th_{aloha-cdma}$) against offered load for different codeset design parameters ($N$, $w$, $\kappa$) for Aloha-CDMA from analysis. The average packet length was 500 bytes and the chipping rate was 1Gb/s.

A graph of throughput against offered load is shown in Figure 3 for the same set of codes analyzed in the previous section. Here load is defined in a similar manner to the previous section. If packets are arriving at a rate $\lambda$ packets per second, the mean number of packets on the line is $\lambda/\mu$ and the offered load on the network is $(\lambda/\mu)(B/N) = \lambda L$ b/s. Expressed as a fraction of the chipping rate this is $\lambda L/B$. An average packet length of 500 bytes and a chipping rate of 1Gb/s were used.

The throughput attains the maximum normalized throughput of around 0.3 at around 50% load. This is the fundamental limit that optical CDMA imposes on the available bandwidth and is independent of the media access mechanism used. As expected, at higher loads we see throughput collapse due to interference. As predicted from the bit collision probabilities, the throughput of a (10, 3, 3) code is marginally higher than that of a (10, 3, 1) code.

At low loads a (100, 3, 3) code performs as well as a (10, 3, 3) code because the sparseness of the 1 chips in the codewords results in low interference. This offsets the increase in the packet service rate. However at higher loads more and more codes are on the line simultaneously due to increased offered load and lower service rate. As a result, the probability of a collision increases and the throughput of the codeset of length 100 degrades to lower than that of the codeset of length 10. Note that throughput only marginally improves by varying the codeset design parameters $N$ and $\kappa$.

## VII. Analysis of *Is/Id* media access

As with Aloha-CDMA, we analyze the *Is/Id* mechanism for bit collision probabilities and network throughput. In an *Is/Id* network, a packet may be lost due to two reasons:

- It may be within the interference sensing limits, but may still cause interference.
- It may be lost because the finite propagation delay of the network prevents perfect interference sensing.

### *A. Interference not prevented by the interference sensing limits*

If the network is always backlogged, then the number of packets on the line will be the maximum possible given the threshold constraints imposed by the interference sensing limits. At any given time let the number of backlogged packets be $l$. Of the $l$ packets offered for transmission, only $m$ packets will actually be transmitted on the line because the *Is/Id* media access mechanism restricts transmission of the packets. We calculate the probability of error in two steps: we calculate the probability of having $m$ packets on the line given that $l$ packets are offered for transmission and then we calculate the probability of error given that $m$ packets are on the line.

The probability of having $m$ packets on the line, given that $l$ packets are offered to the network is

$$P_{line}(m, l) = (1 - P_{thresh}(m)) \prod_{n=m+1}^{l} (P_{thresh}(n))$$

where $P_{thresh}(m)$ is the probability that the threshold ($thresh_c$ overlaps) has been exceeded with $m$ packets on the line.

$$P_{thresh}(m) = \sum_{q=thresh_c}^{N} P_{over}(m, q)$$

where $P_{over}(m, q)$ is the probability of $q$ overlaps, given that $m$ packets are on the line. This is calculated in Appendix III.

The probability of a bit collision when $l$ packets are offered, $P_{error-sensing}(l)$, is given by

$$P_{error-sensing}(l) = \sum_{m=0}^{l} P_{line}(m, l).P_{ber}(m)$$

where $P_{ber}(m)$ is the probability of an error given that $m$ packets are on the line. It is calculated in Appendix IV.

### *B. Interference not sensed due to finite propagation delay*

We also need to take into account the probability of interference that is not sensed due to the finite propagation delay. Two nodes may sense no interference and transmit at the same time and their packets may interfere and be lost. The probability of a packet being corrupted and lost due to interference depends on:

- The number of other packets that arrive/leave during the time when this packet is on the line, i.e., the service time of this packet $t_{service} = 1/\mu = L/(B/N)$
- The number of other packets that arrive during this packet's vulnerable period $t_{vulnerable}$

where $\mu$ is the average service rate, $L$ is the average packet length, $B$ is the overall network bandwidth (the chipping rate) and $N$ is the code length. The probability of a packet collision when $l$ packets are offered to the network is

$$P_{error-collision}(l) = \sum_{m=0}^{l} P_{line}(m, l).P(E_{error}(m))$$

where $E_{error}(m)$ is the event that an error occurs when $m$ packets are on the line during the service time $t_{service}$ of the packet.

Though $m$ packets arrived during the service time, only a few of them, say $q$ would have arrived during the vulnerable period. Therefore,

$$P(E_{error}(m)) = \sum_{q=0}^{m} P(E_{vulnerable}(q,m)).P_{ber}(q)$$

where $E_{vulnerable}(q,m)$ is the event that $q$ packets arrive during the vulnerable period, given that $m$ packets are on the line

The arrival times of packets on the line are controlled by the media access control protocol. However if we assume that the arrival times of packets on the line are uniformly distributed across $t_{service}$ then the probability distribution of the event that $q$ packets arrive during $t_{vulnerable}$, given that $m$ arrivals have occurred during $t_{service}$, is given by

$$P(E_{vulnerable}(q,m)) = \binom{m}{q} p^q (1-p)^{m-q}$$

where $p = t_{vulnerable}/t_{service}$

*C. Overall throughput*

For a packet to be transmitted without error, it must survive both the possible causes of error. Therefore throughput efficiency of the network when $l$ packets are offered for transmission is given by

$$Th_{isid} = (1 - P_{error-collision}(l))(1 - P_{error-sensing}(l))$$

A graph of $Th_{isid}$ against offered load is shown in Figure 4. Note that as the offered load increases, the number of packets on the line increases till the interference sensing limit $thresh_c$ is exceeded. Any increase in offered load after that does not result in more packets on the line. As a result, the bit collision probability remains constant. So does the probability of the threshold failing to catch a case of interference. The parameters used were the same as described in the previous section. The vulnerable time was set to 10$\mu$s, while the thresholds used were $thresh_m = 2$ for both codesets, and $thresh_c = 4$ for the (10, 3, 3) codeset and $thresh_c = 40$ for the (100, 3, 3) codeset. The throughput remains constant above around 30% load.

Note that the network operates at close to optimal throughput after the interference sensing limit is reached even at high loads. In this analysis the only cause of throughput degradation was interference. At high loads, packets are delayed and may time out. We analyze this in the next section.

## VIII. Realistic traffic model

To evaluate the performance of Aloha-CDMA and *Is/Id* on a real network, we modeled realistic traffic patterns on a network.

The traffic model used was based on data obtained from a real LAN. A structural modeling method [27] was used to generate the actual traces used. The traffic trace was generated by a simulation that modeled Internet web traffic characteristics. Packet sizes varied between 40 and 1500 bytes. The generated traffic had an average offered load of around 50Mb/s over a period of 360s. Several traces were merged appropriately to generate higher loads. The network simulated used a 1Gb/s chipping rate and had 100 nodes on the network. Codewords were allocated based on destination addresses. Once the traffic traces were generated, they were fed to a discrete event simulator capable of simulating a network using Aloha-CDMA and *Is/Id*. To determine if a packet was lost due to interference the packet arrival time, its codeword, the other codewords on the line and their relative chip offsets (depending on their arrival times) were used. The simulator was instrumented to measure several parameters: overall throughput, average number of packets on the line, number of transmission retries, number of packets lost due to interference and due to timeouts.

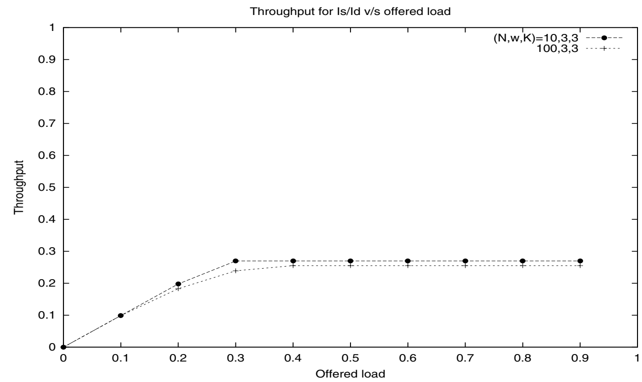


Fig. 4. Throughput for the *Is/Id* mechanism ($Th_{isid}$) against offered load. The average packet length was 500 bytes and the chipping rate was 1Gb/s. The thresholds used were $thresh_m = 2$ for both codesets and $thresh_c$=4 for the (10, 3, 3) codeset and $thresh_c = 40$ for the (100, 3, 3) codeset

The results in Figure 5 show the throughput for the Aloha-CDMA and the *Is/Id* mechanisms. The results are indicated for a (10, 3, 3) codeset. The throughput represents the overall throughput (packets are lost due to both interference and timeouts). The differences in the results when compared to the analysis can be attributed to the bursty nature of real traffic. The network experiences higher loss of packets during periods of burstiness, resulting in higher overall packet collision probabilities. The graphs are shown for different values of backoff count *bc* (100, 500 retries), backoff timer *bt* ($5\mu s$) and $thresh_c$ (4, 6 overlaps). $thresh_m$ was fixed at 2 overlaps.

A graph of the delay against offered load is show in Figure 6. Although the delay rises (and varies a lot as indicated by the standard deviation), if the backoff counter is set fairly low (100 retries) the number of backlogged packets remains stable and the delays remain low. Despite setting the backoff counter low, the throughput doesn't degrade noticeably from its maximum.

Overall, the results indicate that the *Is/Id* mechanism can result in higher throughput on an optical CDMA network. Throughput reductions are due primarily to packets timing out (because the backoff counter has been exceeded) as indicated in Figure 7 which shows the fraction of the total number of packets that are lost due to interference and the fraction lost due to timing out. As can be seen the fraction lost due to

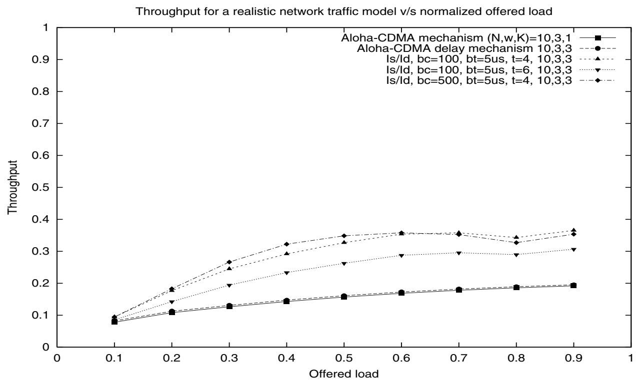


Fig. 5. Throughput against offered load using realistic network traces for the Aloha-CDMA mechanism and *Is/Id*. Plots are shown for different codesets (for Aloha-CDMA) and different values of the backoff timer($bt$), backoff count($bc$) and overlap threshold($t$) (for *Is/Id*).

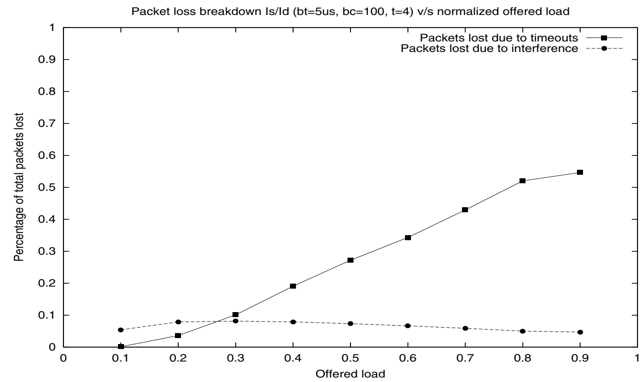


Fig. 7. Breakdown of the fraction of packets dropped (as a fraction of the total packets sent) against offered load using realistic network traces for the *Is/Id* mechanism and a (10, 3, 3) code.

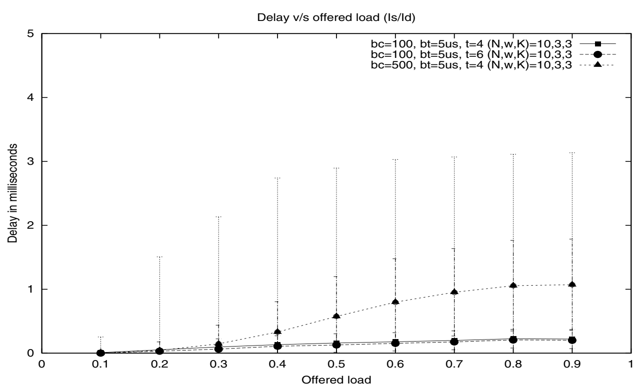


Fig. 6. Delay against offered load using realistic network traces for the *Is/Id* mechanism. Plots are shown for different values of the backoff timer($bt$), backoff count($bc$) and overlap threshold($t$).

interference remains constant.

## IX. Limitations

Our analysis has several limitations. We assumed that multiuser interference is the chief source of error and have neglected other sources of noise such as shot noise and beat noise [4].

We have not discussed how a node will tune its transmitter to the receiver's codeword. We assume that a node may be equipped with multiple decoders and encoders. A fast tunable transmitter/receiver is not a hard requirement of our system, although its presence would make the system more flexible. Recently it has been proposed to use optical microresonators [28] as optical CDMA encoders. These devices, which can be tuned electronically at speeds up to 10GHz, will enable fast tuning of transmitters and receivers.

## X. Conclusion

Optical CDMA networks have been studied for several years. However concerns about their throughput have led to skepticism about their utility.

We proposed the use of *Interference Avoidance* as a method of improving throughput under heavy load. Comparing Aloha-CDMA and the *Is/Id* mechanisms on an optical CDMA link is analogous to the comparison between unslotted Aloha and CSMA/CD mechanisms. Like unslotted Aloha, Aloha-CDMA has zero media access delay and low throughput. The *Is/Id* mechanism improves throughput at the cost of increased delay in a manner similar to CSMA/CD. We analyzed *Is/Id* and have shown it is possible to operate an optical CDMA LAN at close to its maximum possible throughput at high loads. A judicious choice of the interference sensing parameters can ensure that the delay is kept within reasonable bounds. We show that without using interference avoidance, varying the codeset design parameters does not significantly improve the throughput.

An area for further research is a deeper study of interference sensing parameters and how their choice can impact throughput and delay. Another area of research is whether interference avoidance can be used with multiwavelength optical CDMA networks [29].

## References


[1] P. Prucnal, M. Santoro, and T. Fan, "Asynchronous multiplexing for an optical fiber local area network," *IEEE Journal of Lightwave Technology*, vol. 4, p. 547, 1986.

[2] J. Salehi, "Code division multiple-access techniques in optical fiber networks - Part 1: Fundamental principles," *IEEE Transactions on Communications*, vol. 37, no. 8, pp. 824–833, Aug. 1989.

[3] J. Salehi and C. Brackett, "Code division multiple-access techniques in optical fiber networks - Part 2: Systems performance analysis," *IEEE Transactions on Communications*, vol. 37, no. 8, pp. 834–842, Aug. 1989.

[4] C. Lam, "To spread or not to spread: The myths of optical CDMA," in *IEEE Lasers and Electro-Optics Society Annual Meeting*, vol. 2, 2000, pp. 810–811.

[5] H. Chung and P. Kumar, "Optical orthogonal codes - new bounds and an optimal construction," *IEEE Transactions on Information theory*, vol. 36, no. 4, pp. 866–873, July 1990.

[6] S. Johnson, "A new upper bound for error correcting codes," *IRE Transactions on Information Theory*, pp. 203–207, April 1962.

[7] H. Chung, J. Salehi, and V. Wei, "Optical orthogonal codes: Design, analysis, and applications," *IEEE Transactions on Information theory*, vol. 35, no. 3, pp. 595–605, May 1989.

[8] O. Moreno, Z. Zhang, P. Kumar, and V. Zinoviev, "New constructions of optimal cyclically permutable constant weight codes," *IEEE Transactions on Information Theory*, vol. 41, no. 2, pp. 448–455, March 1995.

[9] R. Fuji-Hara and Y. Miao, "Optical orthogonal codes: Their bounds and new optimal constructions," *IEEE Transactions on Information theory*, vol. 46, no. 7, pp. 2396–2406, Nov. 2000.

[10] M. Azizoglu, J. Salehi, and Y. Li, "Optical CDMA via temporal codes," *IEEE Transactions on Communications*, vol. 40, no. 7, pp. 1162–1170, Aug. 1992.

[11] H. Shalaby, "The capacities of optical CDMA communication channels with different code-correlation constraints," in *IEEE Sixth International Symposium on Spread Spectrum Techniques and Applications*, vol. 1, Sep 2000, pp. 179–183.

[12] H. M. H. Shalaby, "Complexities, error probabilites and capacities of optical OOK-CDMA communication systems," *IEEE Transactions on Communications*, vol. 50, no. 12, Dec 2002.

[13] C. S. Hsu and V. O. K. Li, "Performance analysis of slotted fiber-optic code division multiple access (CDMA) packet networks," *IEEE Transactions on Communications*, vol. 45, no. 7, pp. 819 – 828, July 1997.

[14] C. S. Hsu and V. K. Li, "Performance analysis of unslotted fiber-optic code division multiple access (CDMA) packet networks," *IEEE Transactions on Communications*, vol. 45, no. 8, pp. 978 – 987, August 1997.

[15] S. Lee and D. Green, "Performance analysis of optical orthogonal codes in CDMA LANs," *IEE Proceedings*, vol. 145, no. 4, pp. 265–271, August 1998.

[16] J. Muckenheim and K. Iversen, "A new design method for optical CDMA computer networks," in *IEEE ISCC*, 1997.

[17] J. Muckenheim, K. Iversen, and D. Hampicke, "Construction of high-efficient optical CDMA computer networks: Statistical design," in *IEEE International Conference on Communications*, vol. 3, June 1998, pp. 1289–1293.

[18] A. Kumar and D. Patil, "Stability and throughput analysis of CDMA-ALOHA with finite number of users and code sharing," *Telecommunication Systems (a Baltzer Science Journal)*, vol. 8, pp. 257–275, 1997.

[19] D. Raychaudhuri, "Performance analysis of random access packet-switched code division multiple access systems," *IEEE Transactions on Communications*, vol. 29, no. 6, pp. 895– 901, Jun 1981.

[20] A. Polydoros and J. Silvester, "Slotted random access spread-spectrum networks: An analytical framework," *IEEE Journal on Selected Areas in Communications*, vol. 5, no. 6, July 1987.

[21] C. Chae, E. Wong, and R. Tuckker, "Ethernet over passive optical network based on optical CSMA/CD media access technique," in *International Symposium on Services and Local access*, April 2002.

[22] F. Tobagi and V. Hunt, "Performance analysis of carrier sense multiple access with collision detection," *Computer Networks*, vol. 4, no. 5, pp. 245–259, Oct 1980.

[23] M. Molle, "A new binary logarithmic arbitration method for Ethernet," *Technical report CSRI-298, Computer Systems Research Institute, Toronto, Ontario, Canada, 1994.*, 1994.

[24] P. Kamath, J. D. Touch, and J. A. Bannister, "Algorithms for interference sensing in optical CDMA networks," *ISI Technical Report ISI-TR-2003-583*, Nov 2003.

[25] A. Papoulis, "Probability, random variables and stochastic processes," p. 60, 1995.

[26] L. Kleinrock, "Queueing systems Volume I: Theory," pp. 101–102, 1974.

[27] P. Kamath, K. Lan, J. Heidemann, J. Bannister, and J. Touch, "Generation of high bandwidth network traffic traces," in *10th IEEE/ACM International Symposium on Modeling, Analysis and Simulation of Computer Systems (MASCOTS), Fort Worth, Texas*, Oct 2002, pp. 401–410.

[28] K. Djordjev, S. J. Choi, S. J. Choi, and P. D. Dapkus, "Active semiconductor microdisk devices," *IEEE Journal of Lightwave Technology*, vol. 20, no. 1, pp. 105–113, January 2002.

[29] L. Tancevski and I. Andonovic, "Hybrid wavelength hopping/time spreading schemes for use in massive optical networks with increased security," *Journal of Lightwave Technology*, vol. 14, no. 12, pp. 2636–2647, Dec 1996.

## APPENDIX I
## ALOHA-CDMA MECHANISM: PROBABILITY OF BIT COLLISION FOR $\kappa = w$

When $\kappa = w$ there is no constraint on the placement of 1's in a codeword. 1's may be placed anywhere in the codeword. Therefore, for a constant weight codeset of length $N$, weight $w$, if any two codewords are chosen randomly from the sample space consisting of the codewords and their cyclic shifts, then the probability of them having $b$ chips overlapping is given by the *hypergeometric* distribution [25]:

$$P(b) = \frac{\binom{w}{b}\binom{N-w}{w-b}}{\binom{N}{w}}$$

As the number of codewords on the line increases, the probability of an error increases. Consider a receiver tuned to a particular codeword. Let there be $m$ codewords on the line in addition to the codeword being received. These codewords could be any of the codewords belonging to the codeset or their cyclic shifts. We assume that the codewords and their chip offsets are uniformly randomly chosen from the sample space. Let $P_m(b)$ be the probability that $b$ chips overlap between the $m$ codewords and the received codeword.

Then

$$P_m(b) = \sum_{\substack{\forall i,j \\ s.t.\ i+j=b}} P(E_1(i).E_2(j))$$

where $E_1(i)$ is the event that there are $i$ unique overlaps between the $m^{th}$ codeword and the received codeword. $E_2(j)$ is the event that there are $j$ unique overlaps between the $m-1$ preceding codewords and the received codeword.

Then,

$$\begin{aligned} P_m(b) &= \sum_{\substack{\forall i,j \\ s.t.\ i+j=b}} P(E_1(i).E_2(j)) \\ &= \sum_{\substack{\forall i,j \\ s.t.\ i+j=b}} P(E_1(i)|E_2(j)).P(E_2(j)) \\ &= \sum_{\substack{\forall i,j \\ s.t.\ i+j=b}} P(E_1(i)|E_2(j)).P_{m-1}(j) \end{aligned}$$

There have already been $j$ overlaps between the preceding $m-1$ codewords and the codeword to be received. Hence the probability of $i$ new overlaps is the same as the probability of $i$ overlaps in a codeword of weight $w-j$, given $w$ choices. Thus,

$$P(E_1(i)|E_2(j)) = \frac{\binom{(w-j)}{i}\binom{N-(w-j)}{w-i}}{\binom{N}{w}}$$

and from ( 1)

$$P_1(b) = \frac{\binom{w}{b}\binom{N-w}{w-b}}{\binom{N}{w}}$$

An error occurs when the number of overlaps is equal to the receiver threshold. The threshold is generally set to weight $w$ of the code. Therefore, the probability of an error with $m$ codewords on the line is given by $P_m(w)$.

## Appendix II
## Aloha-CDMA mechanism: Probability of bit collision for $\kappa = 1$

The maximum cross correlation between pairs of codes is 1. Therefore, if any two codewords are chosen randomly from the sample space consisting of the codewords and their cyclic shifts, then there are three cases: either 1 chip may overlap, 0 chips may overlap or $w$ chips may overlap (if the same codeword is chosen). The probability of overlap is given by

$$\begin{aligned} P(1) &= w^2/N \\ P(w) &= 1/SN \\ P(0) &= 1 - P(1) - P(w) \\ P(k) &= 0 \qquad k \notin (0, 1, w) \end{aligned}$$

where $S$ is the size of the codeset. We use the the Johnson bound as the size of the codeset. As in Appendix I we can define $P_m(b)$, $E_1$ and $E_2$. The relationship between $P_m(b)$, $E_1$ and $E_2$ follows from Appendix I,

$$P_m(b) = \sum_{\substack{\forall i,j \\ s.t.\ i+j=b}} P(E_1(i)|E_2(j)).P_{m-1}(j)$$

Consider $P(E_1(i)|E_2(j))$. As before, there are three cases to consider. There have already been $j$ overlaps between the preceding $m-1$ codewords and the codeword to be received. Therefore the probability of $i$ new overlaps is the same as the probability of $i$ overlaps in a codeword of weight $w-j$, given $w$ choices.

$$\begin{aligned} P(E_1(1)|E_2(j)) &= \begin{cases} w(w-j)/N \\ \quad when \ \ w-j \neq 1 \\ (w(w-j)S+1)/SN \\ \quad when \ \ w-j = 1 \end{cases} \\ P(E_1(w-j)|E_2(j)) &= 1/SN \\ & \qquad w-j \neq 1 \\ P(E_1(0)|E_j) &= 1 - P(E_1(1)|E_2(j)) \\ & \quad -P(E_1(w-j)|E_2(j)) \\ P(E_1(k)|E_2(j)) &= 0 \\ & \qquad k \notin (0, 1, w-j) \end{aligned}$$

and $P_1(b)$ is given by

$$\begin{aligned} P_1(1) &= w^2/N \\ P_1(w) &= 1/SN \\ P_1(0) &= 1 - P(1) - P(w) \\ P_1(k) &= 0 \qquad k \notin (0, 1, w) \end{aligned}$$

The probability of an error with $m$ codewords on the line is $P_m(w)$.

## Appendix III
## Is/Id mechanism: Probability of an overlap

We assume that the overlap magnitude threshold, $thresh_m$, is fixed at 2. Due to space limitations we provide the analysis only for $\kappa = w$. Let $P_{over}(m, q)$ be the probability that there are a total of $q$ overlapping chip pairs in the $m$ codewords. We denote the hypergeometric probability of choosing $k$ as $hyp(N, K, n, k)$, where $N$ is the total number of possible choices, $K$ the number of favorable choices (the weight), $n$ is the number of trials. Then,

$$P_{over}(m, q) = \sum_{\substack{\forall i,j \\ s.t.\ i+j=q}} P(E_3(i).E_4(j))$$

where $E_3(i)$ is the event that there are $i$ overlaps of magnitude $thresh_m$ between the $m^{th}$ codeword and the preceding $m-1$ codewords.
$E_4(j)$ is the event that there are $j$ overlaps of magnitude $thresh_m$ between the the $m-1$ preceding codewords. Then,

$$\begin{aligned} P_{over}(m, q) &= \sum_{\substack{\forall i,j \\ s.t.\ i+j=q}} P(E_3(i).E_4(j)) \\ &= \sum_{\substack{\forall i,j \\ s.t.\ i+j=q}} P(E_3(i)|E_4(j)).P(E_4(j)) \\ &= \sum_{\substack{\forall i,j \\ s.t.\ i+j=q}} P(E_3(i)|E_4(j)).P_{over}(m-1, j) \end{aligned}$$

When there have already been $j$ overlaps of magnitude 2, the number of possible overlaps remaining is $mw - 2j$. The number of remaining possible overlaps depends on the threshold $thresh_c$. Therefore,

$$P(E_3(i)|E_4(j)) = \begin{cases} hyp(N-j, (mw-2j), w, i) \\ \quad when \ \ w \geq thresh_c - j \\ hyp(N-j, (mw-2j), thresh_c - j, i) \\ \quad when \ \ w < thresh_c - j \end{cases}$$

and,

$$P_{over}(1, q) = hyp(N, w, w, q)$$

## Appendix IV
## Is/Id mechanism: Probability of a bit collision

The analysis for the bit collision probability $P_{ber}(m)$ is similar to the analysis in Appendix I, except that the number of possible overlaps depends on the threshold $thresh_c$. Let $P_m(b)$ be the probability that $b$ chips overlap between the $m$ codewords and the received codeword.

Then,

$$P_m(b) = \sum_{\substack{\forall i,j \\ s.t.\ i+j=b}} P(E_1(i).E_2(j))$$

where $E_1(i)$ is the event that there are $i$ unique overlaps between the $m$th codeword and the received codeword. $E_2(j)$ is the event that there are $j$ unique overlaps between the $m-1$ preceding codewords and the received codeword.

Thus, the probability of b collisions given that $m$ bits are on the line is given by,

$$\begin{aligned} P_m(b) &= \sum_{\substack{\forall i,j \\ s.t.\ i+j=b}} P(E_1(i).E_2(j)) \\ &= \sum_{\substack{\forall i,j \\ s.t.\ i+j=b}} P(E_1(i)|E_2(j)).P(E_2(j)) \\ &= \sum_{\substack{\forall i,j \\ s.t.\ i+j=b}} P(E_1(i)|E_2(j)).P_{m-1}(j) \end{aligned}$$

where,

$$P(E_1(i)|E_2(j)) = \begin{cases} hyp(N-j,(w-j),w,i) \\ \quad when\ \ w \geq thresh_c - j \\ hyp(N-j,(w-j),thresh_c - j, i) \\ \quad when\ \ w < thresh_c - j \end{cases}$$

and,

$$P_1(b) = hyp(N,w,w,b)$$

The bit collision probability given that there are $m$ codewords on the line is $P_{ber}(m) = P_m(w)$.