

# Optimal Memoryless Encoding for Low Power Off-Chip Data Buses

**Yeow Meng Chee\***

School of Physical & Mathematical  
Sciences  
Nanyang Technological University  
Singapore 637616  
ymchee@alumni.uwaterloo.ca

**Charles J. Colbourn**

Department of Computer Science &  
Engineering  
Arizona State University  
Tempe, Arizona 85287-8809  
charles.colbourn@asu.edu

**Alan C. H. Ling**

Department of Computer Science  
University of Vermont  
Burlington, Vermont 05405  
aling@emba.uvm.edu

## Abstract

Off-chip buses account for a significant portion of the total system power consumed in embedded systems. Bus encoding schemes have been proposed to minimize power dissipation, but none has been demonstrated to be optimal with respect to any measure. In this paper, we give the first provably optimal and explicit (polynomial-time constructible) families of memoryless codes for minimizing bit transitions in off-chip buses. Our results imply that having access to a clock does not make a memoryless encoding scheme that minimizes bit transitions more powerful.

## 1. INTRODUCTION

Energy efficiency is an important product quality characteristic. For mobile applications, such as handheld and wireless devices, it not only impacts the usability and acceptance directly, but also affects reliability and packaging cost of the product. Consequently, design techniques for minimizing system power consumption are important for achieving high product quality.

Power-efficient design requires the reduction of power dissipation throughout the design, during all stages of the design process, subject to constraints on system performance and quality of service. In CMOS circuits, most power is dissipated as dynamic power for charging and discharging of internal node capacitances. Thus, researchers have investigated techniques for minimizing the number of transitions inside the circuits. The power dissipation at the input/output (I/O) pads of an integrated circuit (IC) is even higher because off-chip buses have switching capacitances that are orders of magnitude greater than those internal to a chip. The power dissipated at the I/O pads of an IC ranges from 10% to 80% of the total power dissipation with a typ-

\*Y. M. Chee is also with Card View Pte. Ltd., 41 Science Park Road, #04-08A The Gemini, Singapore Science Park II, Singapore 117610.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. ICCAD'06, November 5-9, 2006, San Jose, CA Copyright 2006 ACM 1-59593-389-1/06/0011...\$5.00

ical value of 50% for circuits optimized for low power [18]. The concern of this paper is with low power encoding for off-chip buses.

Bus encoding is used to reduce power dissipated on the bus lines. It has been shown to be an effective technique for power reduction. Compared to on-chip buses of deep submicron (DSM) circuits, the ratio of inter-wire capacitances (or crosstalk coupling) to substrate capacitance is much lower in off-chip buses. The energy model for off-chip buses is thus simpler. We follow the model discussed by Catthoor *et al.* [4]. For a bus with  $k$  wires, each one having metal interconnect capacitance of  $C$ , and under a supply voltage  $V_{dd}$ , the total energy consumed by the bus for a computation of  $N$  cycles can be estimated as

$$E_{bus} = N \cdot k \cdot C \cdot V_{dd}^2. \quad (1)$$

The average power can then be obtained by multiplying  $E_{bus}$  by  $f$ , the frequency at which the bus operates. However, the derivation of (1) assumes that for each cycle, all  $k$  bus wires exhibit a transition that causes the corresponding capacitance to switch. In general, not all the wires switch; hence a cycle-based model is more accurate, as in:

$$E_{bus} = \left( \sum_{i=1}^N k_i \right) \cdot C \cdot V_{dd}^2, \quad (2)$$

where  $k_i$  is the number of wires switching at cycle  $i$ .

The only freedom in (2) for reducing energy consumption is  $k_i$ , since  $N$  depends on the application running on the processor, and  $C$  and  $V_{dd}$  depend on the technology. Therefore, bus encoding techniques for reducing energy consumption have focused on reducing  $k_i$ , the bit switching activity of the off-chip bus. However, no explicit encoding scheme with provable optimality (with respect to any formal measure) is known. In this paper, we give the first explicit and provably optimal memoryless encoding schemes (with respect to rate and maximum energy consumed per cycle), both stateless and stateful, for off-chip buses.

## 2. MATHEMATICAL PRELIMINARIES

### 2.1 Codes

The *Hamming  $n$ -space* is the set  $\mathcal{H}(n) = \{0, 1\}^n$  endowed with the *Hamming distance*  $d_H$  defined as follows: for  $\mathbf{u}, \mathbf{v} \in \mathcal{H}(n)$ ,  $d_H(\mathbf{u}, \mathbf{v})$  is the number of positions where  $\mathbf{u}$  and  $\mathbf{v}$  differ. The *Hamming weight* (for short, *weight*) of a vector  $\mathbf{u} \in \mathcal{H}(n)$  is the number of positions in  $\mathbf{u}$  with non-zero value,

and is denoted  $w_H(\mathbf{u})$ . The  $i$ th component of  $\mathbf{u}$  is denoted  $u_i$ . The *support* of a vector  $\mathbf{u} \in \mathcal{H}(n)$ , denoted  $\text{supp}(\mathbf{u})$ , is the set  $\{i : u_i = 1\}$ .

Any subset of  $\mathcal{H}(n)$  is called a *code of length  $n$* . A *constant weight code of length  $n$  and weight  $w$*  is any subset of  $\mathcal{H}(n, w) = \{\mathbf{u} \in \mathcal{H}(n) : w_H(\mathbf{u}) = w\}$ . The elements of a code are called *codewords*. Let  $\mathcal{C} \subseteq \mathcal{H}(n)$  be a code. The *size* of  $\mathcal{C}$  is  $|\mathcal{C}|$ , the number of codewords in the code. The *diameter* of  $\mathcal{C}$  is  $\text{diam}(\mathcal{C}) = \max_{\mathbf{u}, \mathbf{v} \in \mathcal{C}} d_H(\mathbf{u}, \mathbf{v})$ . Given two codes  $\mathcal{C}_1, \mathcal{C}_2 \subseteq \mathcal{H}(n)$ , the *cross diameter* of  $\mathcal{C}_1$  and  $\mathcal{C}_2$  is  $\text{crossdiam}(\mathcal{C}_1, \mathcal{C}_2) = \max_{\mathbf{u} \in \mathcal{C}_1, \mathbf{v} \in \mathcal{C}_2} d_H(\mathbf{u}, \mathbf{v})$ . By definition,  $\text{crossdiam}(\mathcal{C}, \mathcal{C}) = \text{diam}(\mathcal{C})$ .

## 2.2 Set Systems

For integers  $i < j$ , the set  $\{i, i+1, \dots, j\}$  is abbreviated as  $[i, j]$ . Moreover, we also write  $[j]$  for  $[1, j]$ . For a finite set  $X$  and  $k \leq |X|$ , we define  $2^X = \{A : A \subseteq X\}$  and  $\binom{X}{k} = \{A \in 2^X : |A| = k\}$ . The ring  $\mathbb{Z}/n\mathbb{Z}$  is denoted  $\mathbb{Z}_n$ .

A *set system of order  $n$*  is a pair  $(X, \mathcal{A})$ , where  $X$  is a finite set of  $n$  points and  $\mathcal{A} \subseteq 2^X$ . The elements of  $\mathcal{A}$  are called *blocks*. A set system is said to be  *$k$ -uniform* if  $\mathcal{A} \subseteq \binom{X}{k}$ .

Let  $([n], \mathcal{A})$  be a set system. The *incidence vector* of a block  $A \in \mathcal{A}$  is the vector  $\iota(A) \in \mathcal{H}(n)$ , where

$$\iota(A)_i = \begin{cases} 1, & \text{if } i \in A; \text{ and} \\ 0, & \text{otherwise.} \end{cases}$$

There is a natural correspondence between the Hamming  $n$ -space and the *complete set system of order  $n$* ,  $([n], 2^{[n]})$ : the positions of vectors in  $\mathcal{H}(n)$  correspond to points in  $[n]$ , a vector  $\mathbf{u} \in \mathcal{H}(n)$  corresponds to the block  $\text{supp}(\mathbf{u})$ , and  $d_H(\mathbf{u}, \mathbf{v}) = |\text{supp}(\mathbf{u}) \Delta \text{supp}(\mathbf{v})|$ . From this, it follows that there is a bijection between the set of all codes of length  $n$  and the set of all set systems of order  $n$ . There is also a bijection between the set of all constant weight codes of length  $n$  and weight  $w$  and the set of all  $w$ -uniform set systems of order  $n$ . So we may speak of *the set system of a code*, or *the code of a set system*.

## 3. PROBLEM FORMULATION

Let  $s \geq 1$  and  $n \geq k$ . An  *$s$ -state  $n$ -bit encoding scheme for a source  $S \subseteq \mathcal{H}(k)$*  is a triple  $\mathcal{E} = \langle \mathcal{C}, E, D \rangle$ , where

1.  $\mathcal{C}$  is a code of length  $n$ ,
2.  $E : S \times \mathbb{Z}_s \rightarrow \mathcal{C}$  is an injective map called an *encoding function*, and
3.  $D : \mathcal{C} \times \mathbb{Z}_s \rightarrow S$  is a surjective map called a *decoding function*,

such that  $D(E(\mathbf{u})) = \mathbf{u}$  for all  $\mathbf{u} \in S$ . The encoding function induces a subcode  $\mathcal{C}_i \subseteq \mathcal{C}$  for each  $i \in \mathbb{Z}_s$ , defined by

$$\mathcal{C}_i = \{\mathbf{v} \in \mathcal{C} : \mathbf{v} = E(\mathbf{u}, i) \text{ for some } \mathbf{u} \in S\}.$$

Conversely,  $\{\mathcal{C}_i, i \in \mathbb{Z}_s\}$  uniquely identifies the class of functions that  $E$  (and hence  $D$ ) can belong to:

$$E \in \{f : S \times \mathbb{Z}_s \rightarrow \mathcal{C} : f(\cdot, i) : S \rightarrow \mathcal{C}_i \text{ is injective for all } i \in \mathbb{Z}_s\}.$$

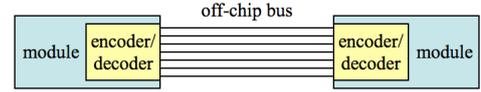


Figure 1: Non-adaptive encoding for off-chip bus

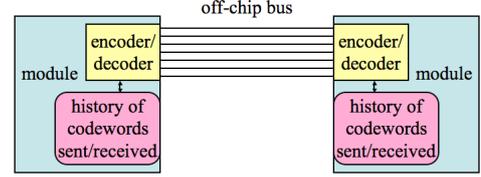


Figure 2: Adaptive encoding for off-chip bus

Encoding and decoding modules that implement  $E$  and  $D$ , respectively, are inserted at the interface of the bus (see Fig. 1). Suppose that  $\mathbf{u}, \mathbf{v} \in S$  are two words to be communicated across the bus at steps  $t$  and  $t+1$  of a computation. In the absence of a bus encoding scheme, the number of bit switchings in this computation cycle is  $|\{i : u_i \neq v_i\}|$ . This quantity is precisely  $d_H(\mathbf{u}, \mathbf{v})$ , which can be as high as  $k$ . If an  $s$ -state  $n$ -bit encoding scheme is used, then  $\mathbf{u}' = E(\mathbf{u}, i)$  and  $\mathbf{v}' = E(\mathbf{v}, j)$  are communicated instead in steps  $t$  and  $t+1$  of the computation, where  $i = t \pmod{s}$  and  $j = t+1 \pmod{s}$ . The resulting number of bit switchings in this computation cycle is therefore  $d_H(\mathbf{u}', \mathbf{v}')$ , which is bounded above by  $\text{crossdiam}(\mathcal{C}_i, \mathcal{C}_j)$ .

To ensure that the encoding scheme guarantees performance under as general a condition as possible, we adopt a worst-case analysis model. Indeed for a given performance guarantee  $\delta \leq k$ , we require that  $\text{crossdiam}(\mathcal{C}_i, \mathcal{C}_{i+1}) \leq \delta$  for all  $i \in \mathbb{Z}_s$ . We call such a family of codes  $(\mathcal{C}_i)_{i \in \mathbb{Z}_s}$  an  $(n, \delta)_s$ -*low power (LP) code*.

A 1-state encoding scheme is known as a *stateless* encoding scheme since it does not need to know the state (step number) of the computation in order to encode the word on the bus. In this case, an  $(n, \delta)_1$ -LP code is a code  $\mathcal{C}$  such that  $\text{diam}(\mathcal{C}) \leq \delta$ . Any  $s$ -state encoding scheme with  $s \geq 2$  is called *stateful*. While a stateless encoding scheme does not need access to a clock, a stateful encoding scheme would require access to a clock to know the particular computation cycle.

The encoding schemes that we consider are *non-adaptive*: The choice of codeword to transmit across the bus in step  $t$  does not depend on codewords that have been transmitted in steps  $i$ ,  $i < t$ . A diagram depicting *adaptive* encoding is given in Fig. 2. Both non-adaptive and adaptive encoding schemes for low-power data buses have been considered (see, for example, [3, 5, 6, 9, 13, 14, 16–20]). Adaptive encoding schemes have better rates. However non-adaptive encoding schemes are often simpler to implement since they do not need to have a continuously changing data model, and do not require memory (to track history of transmitted codewords). Thus non-adaptive encoding schemes are also known as *memoryless* encoding schemes. The focus of this paper is on memoryless encoding schemes.

The *wire expansion* of  $\mathcal{E}$ , an  $n$ -bit encoding scheme for a source  $S \subseteq \mathcal{H}(k)$ , is the quantity  $\alpha(\mathcal{E}) = n/k$ . Wire expansion should be minimized since it affects the system area. Hence, given  $\delta \leq k$ , we want to find an  $(n, \delta)_s$ -LP code  $(\mathcal{C}_i)_{i \in \mathbb{Z}_s}$  with the smallest  $n$  such that  $|\mathcal{C}_i| \geq 2^k$ , for all  $i \in \mathbb{Z}_s$ . In this way we arrive at the following equivalent problem:

**PROBLEM 1.** *Given  $n, \delta$ , and  $s$ , find an  $(n, \delta)_s$ -LP code  $(\mathcal{C}_i)_{i \in \mathbb{Z}_s}$  maximizing  $\min_{i \in \mathbb{Z}_s} |\mathcal{C}_i|$ .*

Let

$$F_s(n, \delta) = \max_{(n, \delta)_s\text{-LP code } (\mathcal{C}_i)_{i \in \mathbb{Z}_s}} \min_{i \in \mathbb{Z}_s} |\mathcal{C}_i|.$$

An  $(n, \delta)_s$ -LP code  $(\mathcal{C}_i)_{i \in \mathbb{Z}_s}$  such that  $\min_{i \in \mathbb{Z}_s} |\mathcal{C}_i| = F_s(n, \delta)$  is called *optimal*.

**LEMMA 1.**  *$F_s(n, \delta) \leq F_2(n, \delta)$  for all  $s \geq 2$ .*

**PROOF.** This follows from the property that if  $(\mathcal{C}_i)_{i \in \mathbb{Z}_s}$  is an  $(n, \delta)_s$ -LP code, then for any  $A \subseteq \mathbb{Z}_s$  such that  $|A| \geq 2$ ,  $(\mathcal{C}_i)_{i \in A}$  is also an  $(n, \delta)_{|A|}$ -LP code.  $\square$

Lemma 1 shows that there is no advantage in using  $s$ -state encoding schemes with  $s > 2$ . Hence, we restrict our attention to  $s = 1$  (the stateless case) and  $s = 2$  (the stateful case) for the rest of this paper.

**PROBLEM 2.** *Given  $n$  and  $\delta$ , construct an optimal  $(n, \delta)_1$ -LP code.*

**PROBLEM 3.** *Given  $n$  and  $\delta$ , construct an optimal  $(n, \delta)_2$ -LP code.*

For  $\mathbf{u}, \mathbf{v} \in \mathcal{H}(n, w)$ , we have  $d_H(\mathbf{u}, \mathbf{v}) \leq \min\{2w, n\}$ . Therefore one way to constrain the number of bit switchings is to limit the code to contain codewords of low weight. This suggests the consideration of constant weight codes.

**PROBLEM 4.** *Given  $n$  and  $\delta$ , construct an optimal  $(n, \delta)_1$ -LP code of constant weight  $w$ .*

**PROBLEM 5.** *Given  $n$  and  $\delta$ , construct an optimal  $(n, \delta)_2$ -LP code of constant weight  $w$ .*

In subsequent sections, we present complete solutions to Problems 2, 3, and 4. Problem 5 is solved for  $n$  sufficiently large.

#### 4. OPTIMAL STATELESS LOW POWER CODES

By definition,  $([n], \mathcal{A})$  is the set system of an  $(n, \delta)_1$ -LP code if and only if  $|A_1 \Delta A_2| \leq \delta$  for all  $A_1, A_2 \in \mathcal{A}$ .

The  $(n, \delta)_1$ -LP codes are known as *anticodes of length  $n$  and maximum distance  $\delta$*  and were introduced in the coding theory literature by Farrell [8] in 1970. However, set systems  $([n], \mathcal{A})$  satisfying  $|A_1 \Delta A_2| \leq \delta$  for all  $A_1, A_2 \in \mathcal{A}$  were studied earlier by Kleitman [12], who obtained a complete solution to the problem of determining the maximum number of blocks in such a set system.

**THEOREM 1 (KLEITMAN).** *Let  $N(n, \delta)$  denote the maximum number of blocks in a set system  $([n], \mathcal{A})$  satisfying  $|A_1 \Delta A_2| \leq \delta$  for all  $A_1, A_2 \in \mathcal{A}$ . Then*

$$N(n, \delta) = \begin{cases} \sum_{i=0}^{\delta/2} \binom{n}{i}, & \text{if } \delta \equiv 0 \pmod{2}; \\ \sum_{i=0}^{(\delta-1)/2} \binom{n}{i} + \binom{n-1}{(\delta-1)/2}, & \text{if } \delta \equiv 1 \pmod{2}. \end{cases}$$

*The extremal set systems are given by*

$$\mathcal{A} = \begin{cases} \bigcup_{i=0}^{\delta/2} \binom{[n]}{i}, & \text{if } \delta \equiv 0 \pmod{2}; \\ \bigcup_{i=0}^{(\delta-1)/2} \binom{[n]}{i} \cup \{A \cup \{x\} : A \in \binom{[n] \setminus \{x\}}{(\delta-1)/2}\}, & \text{if } \delta \equiv 1 \pmod{2}, \end{cases}$$

where  $x$  is any fixed element of  $[n]$ .

The explicit description of optimal  $(n, \delta)_1$ -LP codes provided by Theorem 1 shows that such codes can be constructed in time polynomial in the size of the code. This solves Problem 2.

#### 4.1 Restriction to Constant Weight

We now address constant weight  $(n, \delta)_1$ -LP codes. The diameter of any constant weight  $(n, \delta)_1$ -LP code is even, so we need only consider  $\delta \equiv 0 \pmod{2}$ . In this subsection, we show the equivalence of constant weight  $(n, \delta)_1$ -LP codes and *intersecting families*.

**DEFINITION 1.** *A set system  $(X, \mathcal{A})$  is a  $t$ -intersecting family if  $|A_1 \cap A_2| \geq t$  holds for all  $A_1, A_2 \in \mathcal{A}$ .*

**LEMMA 2.** *The code  $\mathcal{C}$  of a set system  $([n], \mathcal{A})$  is an  $(n, \delta)_1$ -LP code of constant weight  $w$  if and only if  $([n], \mathcal{A})$  is a  $w$ -uniform  $(w - \delta/2)$ -intersecting family.*

**PROOF.** Since  $([n], \mathcal{A})$  is  $(w - \delta/2)$ -intersecting,

$$\begin{aligned} |A_1 \cap A_2| \geq w - \delta/2 &\Leftrightarrow |A_1 \Delta A_2| \leq \delta \\ &\Leftrightarrow d_H(\iota(A_1), \iota(A_2)) \leq \delta, \end{aligned}$$

for all  $A_1, A_2 \in \mathcal{A}$ . Hence,  $\mathcal{C}$  has diameter at most  $\delta$ .  $\square$

Let  $M(n, k, t)$  be the maximum number of blocks in a  $k$ -uniform  $t$ -intersecting family of order  $n$ . The determination of  $M(n, k, t)$  and the structure of maximum  $k$ -uniform  $t$ -intersecting families was initiated by Erdős *et al.* [7], who proved the celebrated Erdős-Ko-Rado Theorem. This result has been improved subsequently by Frankl [10], Wilson [21], and Ahlswede and Khachatrian [2], culminating in the following.

**THEOREM 2 (COMPLETE INTERSECTION THEOREM).**

Let  $1 \leq t \leq k \leq n$  and define  $\mathcal{F}(i) = \{F \in \binom{[n]}{k} : |F \cap [t+2i]| \geq t+i\}$ . Let  $r = \arg \max_{i \in [0, (n-t)/2]} |\mathcal{F}(i)|$ . Then  $M(n, k, t) = |\mathcal{F}(r)|$  and  $\mathcal{F}(r)$  is a maximum  $k$ -uniform  $t$ -intersecting family of order  $n$ .

Lemma 2 together with Theorem 2 completely determines the optimal  $(n, \delta)_1$ -LP codes of constant weight  $w$ , solving Problem 4.

## 5. OPTIMAL STATEFUL LOW POWER CODES

Two set systems  $([n], \mathcal{A})$  and  $([n], \mathcal{B})$  form a  $\delta$ -pair if  $|A \Delta B| \leq \delta$  for all  $A \in \mathcal{A}$  and  $B \in \mathcal{B}$ . Let  $([n], \mathcal{A})$  and  $([n], \mathcal{B})$  be the set systems of  $\mathcal{C}_0$  and  $\mathcal{C}_1$ , respectively. Then from the definition,  $([n], \mathcal{A})$  and  $([n], \mathcal{B})$  form a  $\delta$ -pair if and only if  $(\mathcal{C}_0, \mathcal{C}_1)$  is an  $(n, \delta)_2$ -LP code.

Let  $0 \leq r \leq n$ . A quasi-sphere of dimension  $n$  and radius  $r$  is a set  $\mathcal{Q}$  of the form  $\mathcal{Q} = \{A \subseteq [n] : 0 \leq |A| \leq r\} \cup \bar{\mathcal{Q}}$ , where  $\bar{\mathcal{Q}}$  consists of the first  $L$  elements of  $\{A \subseteq [n] : |A| = r+1\}$  in lexicographic order, for some  $L \geq 0$ . Given  $0 \leq N \leq 2^n$ , there exists a unique quasi-sphere  $\mathcal{Q}$  of dimension  $n$  such that  $|\mathcal{Q}| = N$  (see [11] for a proof).

We want to determine the maximum of  $\min\{|\mathcal{A}|, |\mathcal{B}|\}$ , where  $([n], \mathcal{A})$  and  $([n], \mathcal{B})$  form a  $\delta$ -pair. A related result has been obtained by Ahlswede and Katona [1].

**THEOREM 3 (AHLWEDE AND KATONA).** Let  $1 \leq N \leq 2^n$ ,  $1 \leq \delta \leq n$ , and let  $([n], \mathcal{A})$  and  $([n], \mathcal{B})$  be a  $\delta$ -pair with  $|\mathcal{A}| = N$ . Then  $\max |\mathcal{B}|$  is achieved if  $\mathcal{A}$  is a quasi-sphere and  $\mathcal{B} = \{B : |B \Delta A| \leq \delta \text{ for all } A \in \mathcal{A}\}$ .

We assume without loss of generality that  $\delta < n$ , since otherwise we may take  $\mathcal{A} = \mathcal{B} = 2^{[n]}$ . Suppose that  $([n], \mathcal{A})$  and  $([n], \mathcal{B})$  form a  $\delta$ -pair, with  $|\mathcal{A}| = N$ , such that  $\mathcal{B}$  has the maximum number of blocks. Write  $N = \sum_{i=0}^r \binom{n}{i} + m$ ,  $0 \leq m < \binom{n}{r+1}$ , for some  $r \leq n$ . Theorem 3 asserts that  $\mathcal{A}$  is the quasi-sphere of dimension  $n$  and radius  $r$ , with  $|\mathcal{A}| = N$ , and  $\mathcal{B} = \{B \subseteq [n] : |B \Delta A| \leq \delta \text{ for all } A \in \mathcal{A}\}$ . Suppose that  $B \in \mathcal{B}$  and  $|B| = k$ . Then there exists  $A \in ([n] \setminus B)$  with  $|A| = \delta - k + 1$  whenever  $\delta < n$ . Since  $|A \Delta B| = \delta + 1$ , we have  $A \notin \mathcal{A}$ . This is possible only if  $\delta - k + 1 > r$ , or  $k < \delta - r + 1$ . It follows that no  $k$ -subset of  $[n]$ ,  $k \geq \delta - r + 1$ , is in  $\mathcal{B}$ . Hence,  $|\mathcal{B}| \leq \sum_{i=0}^{\delta-r} \binom{n}{i}$ .

Now consider the case that  $\delta \equiv 0 \pmod{2}$ . If  $r \geq \delta/2$ , then  $|\mathcal{B}| \leq \sum_{i=0}^{\delta/2} \binom{n}{i} \leq N(n, \delta)$ . If  $r < \delta/2$ , then  $|\mathcal{A}| \leq \sum_{i=0}^{\delta/2-1} \binom{n}{i} + \binom{n}{\delta/2} = \sum_{i=0}^{\delta/2} \binom{n}{i} \leq N(n, \delta)$ . Hence, we have  $\min\{|\mathcal{A}|, |\mathcal{B}|\} \leq N(n, \delta)$ .

Next, consider the case that  $\delta \equiv 1 \pmod{2}$ . If  $r \geq (\delta+1)/2$ , then  $|\mathcal{B}| \leq \sum_{i=0}^{(\delta-1)/2} \binom{n}{i} \leq N(n, \delta)$ . If  $r \leq (\delta-3)/2$ , then  $|\mathcal{A}| \leq \sum_{i=0}^{(\delta-3)/2} \binom{n}{i} + \binom{n}{(\delta-1)/2} \leq \sum_{i=0}^{(\delta-3)/2} \binom{n}{i} + \binom{n-1}{(\delta-1)/2} \leq N(n, \delta)$ . Hence,  $\min\{|\mathcal{A}|, |\mathcal{B}|\} \leq N(n, \delta)$ , when  $r \neq (\delta-1)/2$ .

We deal with the remaining case in which  $\delta \equiv 1 \pmod{2}$  and  $r = (\delta-1)/2$ . If  $m \leq \binom{n-1}{r}$ , then  $|\mathcal{A}| \leq \sum_{i=0}^{(\delta-1)/2} \binom{n}{i} +$

$\binom{n-1}{(\delta-1)/2} \leq N(n, \delta)$ . If  $m > \binom{n-1}{r}$ , then by definition of a quasi-sphere,  $\mathcal{A}$  contains the sets  $\{1\} \cup R$ , where  $R \in \binom{[2, n]}{r}$ . We claim that  $1 \in B$  for every  $B \in \mathcal{B}$  such that  $|B| = r+1$ . Suppose not, let  $R \subseteq [n] \setminus B$  such that  $|R| = r$ . Then  $A = \{1\} \cup R \in \mathcal{A}$ , and  $|A \Delta B| = 2r+2 > \delta$ , giving a contradiction. It follows that all subsets of size  $r+1$  in  $\mathcal{B}$  must contain the element 1, and hence the number of subsets of size  $r+1$  in  $\mathcal{B}$  is at most  $\binom{n-1}{r}$ . Hence,  $|\mathcal{B}| \leq \sum_{i=0}^{(\delta-1)/2} \binom{n}{i} + \binom{n-1}{(\delta-1)/2} \leq N(n, \delta)$ .

This establishes the following:

**THEOREM 4.** For any  $\delta$ -pair  $([n], \mathcal{A})$  and  $([n], \mathcal{B})$ ,

$$\min\{|\mathcal{A}|, |\mathcal{B}|\} \leq \begin{cases} \sum_{i=0}^{\delta/2} \binom{n}{i}, & \text{if } \delta \equiv 0 \pmod{2}; \\ \sum_{i=0}^{(\delta-1)/2} \binom{n}{i} + \binom{n-1}{(\delta-1)/2}, & \text{if } \delta \equiv 1 \pmod{2}, \end{cases}$$

with equality if

$$\mathcal{A} = \mathcal{B} = \begin{cases} \bigcup_{i=0}^{\delta/2} \binom{[n]}{i}, & \text{if } \delta \equiv 0 \pmod{2} \\ \bigcup_{i=0}^{(\delta-1)/2} \binom{[n]}{i} \cup \{A \cup \{x\} : A \in \binom{[n] \setminus \{x\}}{(\delta-1)/2}\}, & \text{if } \delta \equiv 1 \pmod{2}, \end{cases}$$

where  $x$  is any fixed element of  $[n]$ .

**COROLLARY 1.**  $F_1(n, \delta) = F_2(n, \delta)$ ; the size of an optimal  $(n, \delta)_2$ -LP code is the same as the size of an optimal  $(n, \delta)_1$ -LP code.

This solves Problem 3.

### 5.1 Restriction to Constant Weight

Two set systems  $([n], \mathcal{A})$  and  $([n], \mathcal{B})$  are called *cross-wise  $t$ -intersecting* if  $|A \cap B| \geq t$  for all  $A \in \mathcal{A}$  and  $B \in \mathcal{B}$ .

**LEMMA 3.** Let  $([n], \mathcal{A})$  and  $([n], \mathcal{B})$  be the set systems of codes  $\mathcal{C}_0$  and  $\mathcal{C}_1$ , respectively. Then  $(\mathcal{C}_0, \mathcal{C}_1)$  is an  $(n, \delta)_2$ -LP code of constant weight  $w$  if and only if  $([n], \mathcal{A})$  and  $([n], \mathcal{B})$  are  $w$ -uniform and cross-wise  $(w - \delta/2)$ -intersecting.

**PROOF.** Follows from the observation that  $([n], \mathcal{A})$  and  $([n], \mathcal{B})$  are cross-wise  $(w - \delta/2)$ -intersecting if and only if for all  $A \in \mathcal{A}$  and  $B \in \mathcal{B}$ ,

$$\begin{aligned} |A \cap B| \geq w - \delta/2 &\Leftrightarrow |A \Delta B| \leq \delta \\ &\Leftrightarrow d_H(\iota(A), \iota(B)) \leq \delta. \end{aligned}$$

Hence,  $(\mathcal{C}_0, \mathcal{C}_1)$  is an  $(n, \delta)_2$ -LP code.  $\square$

Let  $([n], \mathcal{A})$  and  $([n], \mathcal{B})$  be two set systems that are  $w$ -uniform cross-wise  $t$ -intersecting, where  $t = w - \delta/2$ . Suppose that  $A_1, A_2 \in \mathcal{A}$ . Let  $B \in \mathcal{B}$ ,  $U \subseteq A_1 \cap A_2 \cap B$ , and  $u = |U|$ . Then  $B$  must contain a further  $t - u$  points from each of  $A_1 \setminus U$  and  $A_2 \setminus U$ . The remaining  $w - (2t - u)$  points in  $B$  are from the  $n - (2t - u)$  points not already contained in  $B$ . The total possible number of such blocks  $B$  is then

$$\binom{w-u}{t-u}^2 \binom{n-(2t-u)}{w-(2t-u)}. \quad (3)$$

If  $u \leq t - 1$ , then for  $n$  large enough, (3) is less than  $M(n, k, t) = \binom{n-t}{w-t}$ , and hence  $|\mathcal{B}| \leq M(n, k, t)$ .

On the other hand, if  $u \geq t$ , then  $|A_1 \cap A_2| \geq t$ , which implies  $\mathcal{A}$  is  $t$ -intersecting. So we also have  $|\mathcal{A}| \leq M(n, k, t)$ . This gives the following result.

**THEOREM 5.** *Suppose  $([n], \mathcal{A})$  and  $([n], \mathcal{B})$  are  $w$ -uniform cross-wise  $t$ -intersecting set systems. Then for large enough  $n$ ,  $\min\{|\mathcal{A}|, |\mathcal{B}|\} \leq \binom{n-t}{w-t}$ , with equality if  $\mathcal{A} = \mathcal{B} = \{A \in \binom{[n]}{w} : T \subseteq A\}$ , where  $T$  is any fixed  $t$ -subset of  $[n]$ .*

**COROLLARY 2.** *For  $n$  large enough, the size of an optimal  $(n, \delta)_2$ -LP code of constant weight  $w$  is the same as the size of an optimal  $(n, \delta)_1$ -LP code of constant weight  $w$ .*

This solves Problem 5 for  $n$  large enough. A solution for all  $n$  seems out of reach at the moment.

## 6. IMPLEMENTATION

The encoding schemes introduced can be easily implemented with a look-up table, since our codes are explicit and constructible in polynomial time. However, it is also possible to encode and decode algorithmically, removing the need to store a look-up table of size  $2^k$  when the source is  $\mathcal{H}(k)$ . We illustrate this with an encoding/decoding algorithm for the optimal  $(n, \delta)_1$ -LP code from Section 4. We assume that the source is  $\mathcal{H}(k)$ .

First we define ranking and unranking algorithms. Given a set of objects  $X$ , *rank* and *unrank* are functions  $f : X \rightarrow \{0, \dots, |X|-1\}$  and  $g : \{0, \dots, |X|-1\} \rightarrow X$ , such that  $f$  and  $g$  are bijections satisfying  $g(f(x)) = x$  for all  $x \in X$ . Computing  $f$  is *ranking*, and computing  $g$  is *unranking*. Efficient ranking and unranking algorithms for  $\mathcal{H}(n, w)$  (equivalent to  $w$ -subsets of an  $n$ -set) are well known (see, for example, [15]). A simple and efficient ranking and unranking algorithm for  $\mathcal{H}(n, w)$  using the *co-lexicographic* ordering is described below.

<pre> <b>rank</b>(<math>w, \mathbf{x}</math>) {   <math>\{t_1, \dots, t_w\} = \text{supp}(\mathbf{x});</math>   <math>r = 0;</math>   for <math>i = 1</math> to <math>w</math>     <math>r = r + \binom{t_i-1}{k+1-i};</math>   return <math>r;</math> }</pre>	<pre> <b>unrank</b>(<math>n, w, r</math>) {   <math>x = n;</math>   for <math>i = 1</math> to <math>w</math> {     while <math>\binom{x}{w+1-i} &gt; r</math>       <math>x = x - 1;</math>     <math>t_i = x + 1;</math>     <math>r = r - \binom{x}{k+1-i};</math>   }   return <math>\iota(\{t_1, \dots, t_w\});</math> }</pre>
--	--

We now give a high-level overview of the method. We interpret  $\mathbf{x} \in \mathcal{H}(k)$  as an integer in the interval  $[0, 2^k - 1]$  in the natural way. Let  $N_i = \sum_{j=0}^i \binom{n}{j} - 1$ , for  $0 \leq i \leq n$ . By convention,  $N_{-1} = -1$ . An  $\mathbf{x} \in [N_{t-1} + 1, N_t]$  is encoded as an element of  $\mathcal{H}(n, t)$  by unranking  $\mathbf{x} - N_{t-1} - 1$ . To decode a received  $\mathbf{y} \in \mathcal{H}(n)$ , simply rank  $\mathbf{y}$ .

More precisely, to encode a source  $\mathbf{x} \in \mathcal{H}(k)$  to a codeword  $\mathbf{y}$ , the encoder performs the following steps:

**Step 1:** Find  $t$  such that  $N_{t-1} < \mathbf{x} \leq N_t$ .

**Step 2:**  $m = \mathbf{x} - N_{t-1} - 1$ .

**Step 3:**  $\mathbf{y} = \text{unrank}(n, t, m)$ .

The decoding algorithm is even simpler. To decode  $\mathbf{y} \in \mathcal{H}(n)$  to a source vector  $\mathbf{x}$ , the decoder performs the following steps:

**Step 1:** Compute  $w = w_H(\mathbf{y})$ .

**Step 2:**  $\mathbf{x} = \text{rank}(w, \mathbf{y})$ .

An optimal constant weight  $(n, \delta)_1$ -LP code can be similarly implemented.

## 7. NOT ALL LIMITED WEIGHT CODES ARE EQUAL

In [19], Stan and Burleson introduced the following bus encoding scheme. Define an  *$m$ -limited weight code* ( $m$ -LWC) of length  $n$  to be a code  $\mathcal{C} \subseteq \cup_{w=0}^m \mathcal{H}(n, w)$ . Suppose that the number of source states to be transmitted across a bus is  $2^k$  and that the source states are to be encoded with a code of length  $n$ . Let  $m$  be the smallest integer such that

$$\sum_{i=0}^m \binom{n}{i} \geq 2^k.$$

Stan and Burleson claimed that an  $m$ -LWC code of length  $n$ , comprising as codewords all elements in  $\cup_{w=0}^{m-1} \mathcal{H}(n, w)$  and the remaining  $2^k - \sum_{i=0}^m \binom{n}{i}$  codewords from  $\mathcal{H}(n, m)$ , is

*“optimal in the sense that any other code with the same length cannot have better statistical properties for low power.”*

This statement is true since we have shown in Section 4 that an optimal  $(n, \delta)_1$ -LP code is an  $m$ -LWC, but we must exercise caution when choosing the  $m$ -LWC, since not every  $m$ -LWC is optimal, as can be seen in the example below.

**EXAMPLE 1.** *Let  $k = 5$  and  $n = 6$ . We give two 3-LWCs of length  $n$  and size  $2^k$  with different diameters. Each of the two 3-LWCs of length  $n$  and size  $2^k$  contains all 22 elements in  $\cup_{w=0}^2 \mathcal{H}(6, w)$ . The first 3-LWC containing the 10 additional codewords*

```

000111 001011 010011 100011 001110
111000 110100 101100 011100 110001
```

*has diameter six, and the second 3-LWC containing the 10 additional codewords*

```

100011 100101 100110 101001 101010
101100 110001 110010 110100 111000
```

*has diameter five.*

Our results in Section 4 give  $m$ -LWCs that are optimal, for every length.

## 8. CONCLUSION

Past research on encoding schemes for low-power buses has largely been experimental; no optimal codes were known for any measure, in any model. This paper constructs codes that are provably optimal, starting with the simplest model, that of an off-chip data bus. In so doing, we obtain the first explicit and provably optimal memoryless encoding scheme that minimizes bit transitions for off-chip data buses. Our approach is combinatorial and the codes obtained are explicit and polynomial-time constructible. We also show that having access to a clock (or alternatively, knowing the computation cycle) does not help in achieving more efficient encoding.

We are currently extending this work to thermal-aware models and models at the DSM level, where inter-wire capacitances are significant, and crosstalks must be avoided.

## Acknowledgments

The authors are grateful to Rudolf Ahlswede for helpful pointers.

## 9. REFERENCES

- [1] R. Ahlswede and G. O. H. Katona. Contributions to the geometry of Hamming spaces. *Discrete Math.*, 17:1–22, 1977.
- [2] R. Ahlswede and L. H. Khachatrian. The complete intersection theorem for systems of finite sets. *European J. Combin.*, 18(2):125–136, 1997.
- [3] L. Benini, G. de Micheli, E. Macii, D. Sciuto, and C. Silvano. Asymptotic zero-transition activity encoding for address busses in low-power microprocessor-based systems. In *GLSVLSI '97: Proceedings of the 7th Great Lakes Symposium on VLSI*, pages 77–82. IEEE Computer Society, 1997.
- [4] F. Catthoor, S. Wuytack, E. De Greef, F. Balasa, L. Nachtergaele, and A. Vandecappelle. *Custom Memory Management Methodology - Exploration of Memory Organization for Embedded Multimedia System Design*. Kluwer Acad. Publ., Boston, 1998.
- [5] W.-C. Cheng and M. Pedram. Memory bus encoding for low power: A tutorial. In *ISQED 2001: Proceedings of the 2nd International Symposium on Quality of Electronic Design*, pages 199–204. IEEE Computer Society, 2001.
- [6] W.-C. Cheng and M. Pedram. Power-optimal encoding for a DRAM address bus. *IEEE Trans. Very Large Scale Integr. Syst.*, 10(2):109–118, 2002.
- [7] P. Erdős, C. Ko, and R. Rado. Intersection theorems for systems of finite sets. *Quart. J. Math. Oxford Ser. (2)*, 12:313–320, 1961.
- [8] P. G. Farrell. Linear binary anticodes. *Electron. Lett.*, 6:419–421, 1970.
- [9] W. Fornaciari, M. Polentarutti, D. Sciuto, and C. Silvano. Power optimization of system-level address buses based on software profiling. In *CODES '00: Proceedings of the Eighth International Workshop on Hardware/Software Codesign*, pages 29–33. ACM Press, 2000.
- [10] P. Frankl. The Erdős-Ko-Rado theorem is true for  $n = kct$ . In *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. I*, volume 18 of *Colloq. Math. Soc. János Bolyai*, pages 365–375. North-Holland, Amsterdam, 1978.
- [11] G. O. H. Katona. The Hamming-sphere has minimum boundary. *Studia Sci. Math. Hungar.*, 10(1-2):131–140, 1975.
- [12] D. J. Kleitman. On a combinatorial conjecture of Erdős. *J. Combin. Theory*, 1:209–214, 1966.
- [13] T. Lindkvist, J. Lofvenberg, H. Ohlsson, K. Johansson, and L. Wanhammar. A power-efficient, low-complexity, memoryless coding scheme for buses with dominating inter-wire capacitances. In *IWSOC '04: Proceedings of the 4th IEEE International Workshop on System-on-Chip for Real-Time Applications*, pages 257–262. IEEE Computer Society, 2004.
- [14] R. Murgai and M. Fujita. On reducing transitions through data modifications. In *DATE '99: Proceedings of the Conference on Design, Automation and Test in Europe*, pages 82–88. ACM Press, 1999.
- [15] A. Nijenhuis and H. S. Wilf. *Combinatorial Algorithms*. Academic Press, New York, 1978.
- [16] K. N. Patel and I. L. Markov. Error-correction and crosstalk avoidance in DSM busses. *IEEE Trans. Very Large Scale Integr. Syst.*, 12(10):1076–1080, 2004.
- [17] Y. Shin, S.-I. Chae, and K. Choi. Partial bus-invert coding for power optimization of system level bus. In *ISLPED '98: Proceedings of the 1998 International Symposium on Low Power Electronics and Design*, pages 127–129. ACM Press, 1998.
- [18] M. R. Stan and W. P. Burleson. Bus-invert coding for low-power I/O. *IEEE Trans. Very Large Scale Integr. Syst.*, 3(1):49–58, 1995.
- [19] M. R. Stan and W. P. Burleson. Coding a terminated bus for low power. In *GLSVLSI '95: Proceedings of the Fifth Great Lakes Symposium on VLSI*, pages 70–73. IEEE Computer Society, 1995.
- [20] P. Subrahmanya, R. Manimegalai, V. Kamakoti, and M. Mutyam. A bus encoding technique for power and cross-talk minimization. In *VLSI Design 2004: 17th International Conference on VLSI Design*, pages 443–448. IEEE Computer Society, 2004.
- [21] R. M. Wilson. The exact bound in the Erdős-Ko-Rado theorem. *Combinatorica*, 4(2-3):247–257, 1984.