Optimal Memoryless Encoding for Low Power Off-Chip Data Buses

Yeow Meng Chee

School of Physical & Mathematical Sciences
Nanyang Technological University
Singapore

Charles J. Colbourn

Dept of Computer Science & Engineering
Arizona State University
USA

Alan C. H. Ling

Dept of Computer Science
University of Vermont
USA

Motivation

- High switching capacitance of off-chip data buses leads to high power dissipation at I/O pads:
  - 10%–80% of total power dissipated [Stan & Burleson, 1995]
- Bus encoding is known to be an effective technique to reduce power dissipation on data buses:
  - Bus Invert Coding [Stan & Burleson, 1995]
  - Two-Dimensional Coding [Stan & Burleson, 1996]
  - Constant Weight Coding [Tallini & Bose, 1998]
  - Offset-Coding [Fornaciari et al., 2000]
  - Transition Pattern Coding [Sotiriadis & Chadrakasan, 2003]
- No optimal encoding with respect to any formal model is known.

Contributions

- Positive result
  - The first provably optimal encoding scheme for low power buses.
    - Code is explicit: polynomial-time constructible.
    - Simple implementation.
- Negative result
  - Memoryless encoding schemes do not become more powerful when they have access to a clock.

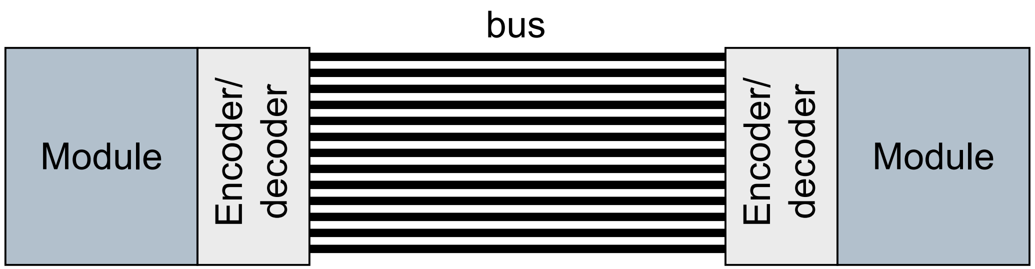
Power model for off-chip data buses

- Energy dissipated by a bus on a computation of N cycles (Catthoor et al., 1998):

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

  - E_{bus} - energy dissipated by the bus.
  - k_i - number of wires switching at cycle i.
  - N - number of computation cycles.
  - C - interconnect capacitance of each wire.
  - V_{dd} - supply voltage.
- Only freedom for reducing power is k_i.
  - N - depends on application running.
  - C, V_{dd} - depends on technology.

Encoding schemes



- To transmit a word u from one module to another, the transmitting module encodes u and sends it across the bus.
- The receiving module decodes the received word to recover u.
- Two categories of encoding schemes:
  - Adaptive
    - Encoding function depends on previously transmitted codewords (necessary to keep history of transmitted codewords).
  - Nonadaptive
    - Encoding function does not depend on previously transmitted codewords (requires no memory to track transmitted codewords).
  - For this reason, nonadaptive is also called memoryless.

Encoding schemes with access to clocks

- An encoding scheme with access to a clock can tell which step of the computation cycle it is in.
  - Also known as a stateful encoding scheme.
- Encoding scheme without access to a clock is a stateless encoding scheme.
- Our focus is on
  - Memoryless stateless encoding schemes
  - Memoryless stateful encoding schemes

Problem formulation

- Hamming k-space: $H(k) = \{0,1\}^k$
- Hamming distance function: $d_H(\cdot,\cdot)$
- Suppose two modules need to communicate source states from $H(k)$ across a bus.
- An stateless encoding scheme is an injection $E : H(k) \to C \subseteq H(n)$ that maps a k-bit word to an n-bit codeword and sends it across the bus.
- Under a worst case analysis model, the number of bit transitions between $E(u)$ and $E(v)$ must be not more than a given guarantee δ, for any source states u and v:

$$d_H(x,y) \leq \delta$$

  for any $x,y \in C$
- Such a C is called an (n,δ)-LP code.
- We want to minimize n, since this is the number of wires needed for the bus (which affects system area).

Problem formulation (cont.)

- Equivalently, given n and δ, we want to find an (n,δ)-LP code of maximum size.

Connections

- An (n,δ)-LP code is known as an anticode to the coding theory community.
- The problem of determining the largest size, $N(n,\delta)$, of an (n,δ)-LP code was raised by Erdös and solved by Katona (1964) and Kleitman (1966):

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

- Idea
  - If $\delta = 2r$, take as codewords all vectors of weight at most r.
  - If $\delta = 2r+1$, take as codewords all vectors of weight at most r, and all codewords of weight $r+1$ having first component one.

Constant weight LP codes

- If u and v are length n vectors of constant weight w, then

$$d_H(u,v) \leq \min\{2w, n\}$$

- This suggests the consideration of constant weight codes to minimize transitions.
- A set system of order n is a collection of subsets of $\{1,2,\ldots,n\}$, called blocks.
- A set system is w-uniform if its blocks are all of size w.
- A set system is t-intersecting if any pair of blocks intersect in at least t elements.
- An (n,δ)-LP code consisting of vectors of constant weight w is equivalent to a w-uniform $(w-\delta/2)$-intersecting set system of order n.

Constant weight LP codes (cont.)

- The problem of determining the maximum size of a w-uniform t-intersecting set system has a long history beginning with the celebrated Erdos-Ko-Rado Theorem in 1961 and culminating in the Complete Intersection Theorem of Ahlswede & Khachatrian in 1997.
- The Complete Intersection Theorem completely solves the problem of determining the maximum size of an (n,δ)-LP code of constant weight w.

Stateful Encoding Schemes

- The problem formulation for stateful encoding schemes are similar.
- Surprisingly, the knowledge of state (access to a clock) does not help.
- We are able to show that, in general, stateful encoding schemes are no more powerful that stateless ones.
- If we restrict to constant weight encoding schemes, then we are able to show that for large enough k, stateful encoding schemes of constant weight w is no more powerful than stateless ones.

Implementation

- How do we implement the encoding and decoding functions?
- Table look-up is unacceptable as it may waste more power than save.
- We give a method for encoding and decoding algorithmically.
- The idea is based on ranking and unranking (with respect to co-lexicographic ordering) the set of vectors in $H(n)$ of constant weight w.

Conclusion

- Past research on low power bus encoding has been empirical.
  - No optimal encoding schemes known with respect to any formal model.
- Our results are the first provably optimal codes for the simplest model of an off-chip data bus.
- Approach is combinatorial and codes obtained are explicit and polynomial-time constructible
- We are currently working on extensions to thermal-aware models and DSM models, where inter-wire capacitances are significant and crosstalks must be avoided.