

# Timing-Aware Power-Optimal Ordering of Signals

KONSTANTIN MOISEEV and AVINOAM KOLODNY

Technion, Haifa, Israel

and

SHMUEL WIMER

Intel Corporation, Haifa, Israel and Technion, Haifa, Israel

---

A computationally efficient technique for reducing interconnect active power in VLSI systems is presented. Power reduction is accomplished by simultaneous wire spacing and net ordering, such that cross-capacitances between wires are optimally shared. The existence of a unique power-optimal wire order within a bundle is proven, and a method to construct this order is derived. The optimal order of wires depends only on the activity factors of the underlying signals; hence, it can be performed prior to spacing optimization. By using this order of wires, optimality of the combined solution is guaranteed (as compared with any other ordering and spacing of the wires). Timing-aware power optimization is enabled by simultaneously considering timing criticality weights and activity factors for the signals. The proposed algorithm has been applied to various interconnect layouts, including wire bundles from high-end microprocessor circuits in 65 nm technology. Interconnect power reduction of 17% on average has been observed in such bundles.

Categories and Subject Descriptors: B.7.2 [**Integrated Circuits**]: Design Aids—*Layout, placement and routing*; J.6 [**Computer-Aided Engineering**]—*Computer-aided design (CAD)*

General Terms: Algorithms, Design

Additional Key Words and Phrases: Wire ordering, wire spacing, power optimization, interconnect optimization

**ACM Reference Format:**

Moiseev, K., Kolodny, A., and Wimer, S. 2008. Timing-aware power-optimal ordering of signals. ACM Trans. Des. Autom. Elect. Syst. 13, 4, Article 65 (September 2008), 17 pages. DOI = 10.1145/1391962.1391973 <http://doi.acm.org/10.1145/1391962.1391973>

---

Authors' addresses: K. Moiseev and A. Kolodny, Department of Electrical Engineering, Technion—Israel Institute of Technology, 33000, Haifa, Israel, E-mail: {mkostya,kolodny}@ee.technion.ac.il; S. Wimer, Intel Corporation, Israel Development Center, 31015, Haifa, Israel, E-mail: shmuel.wimer@intel.com.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or direct commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701 USA, fax +1 (212) 869-0481, or permissions@acm.org. © 2008 ACM 1084-4309/2008/09-ART65 \$5.00 DOI 10.1145/1391962.1391973 <http://doi.acm.org/10.1145/1391962.1391973>

ACM Transactions on Design Automation of Electronic Systems, Vol. 13, No. 4, Article 65, Pub. date: Sept. 2008.

## 1. INTRODUCTION

With the advancement of semiconductor technology, power dissipation becomes an important design objective. Interconnect power, which is the power dissipated by charging and discharging of wire capacitances, typically represents about 50% of the circuit's dynamic power [Magen et al. 2004]. Therefore, the optimization of interconnect power is an important VLSI design challenge. Interconnect power can be expressed by

$$P_{tot} \propto \sum_{i=1}^N C_i \cdot V_{DD} \cdot V_i \cdot f \cdot AF_i, \quad (1.1)$$

where summation is done over all  $N$  nodes of the circuit,  $C_i$  is the interconnect capacitance at node  $i$ ,  $V_i$  is the voltage swing at node  $i$ ,  $f$  is the clock frequency and  $AF_i$  is the activity factor of node  $i$ . Common power reduction techniques are based on architectural, logic or circuit design methods, decreasing  $f$ ,  $AF_i$ ,  $N$  or  $V_i$  [Zang et al. 2000; Hsieh and Pedram 2000; Shang et al. 2002]. Bus encoding techniques for reducing activity and wire cross-capacitance have also been used [Bertozzi et al. 2002; Benini et al. 1998; Kim et al. 2000; Cha et al. 2006].

In this article, we propose a technique for reduction of interconnect power by reducing the capacitance term  $C_i$  in (1.1) for the most active nodes within parallel wire bundles. Wire bundles are common in global interconnect structures of modern VLSI circuits. The capacitances in such structures (see Figure 1) are typically dominated by cross-capacitances between adjacent wires, since the aspect ratio of wire thickness to wire spacing grows with the progression of manufacturing technology [ITRS 2005].

The proposed method finds the best permutation of signals in the bundle and sets the physical positions of the wires, such that the most active wires will share the smallest cross-capacitances. Unlike encoding, which works in the logic domain, adding special logic that consumes area and increases delay, the method proposed in this paper works in the physical layout domain and doesn't add any logic and area. The method is extended to consider timing-critical signals in determination of the wire order and inter-wire spaces.

Optimal cross-capacitance sharing for power reduction is achieved by wire reordering and space allocation according to activity factors of the signals. Signals with high activity should be loaded by small cross-capacitances, which are obtained by large spaces. Low-activity signals can tolerate smaller spaces. In order to best utilize the given area, which is a constrained resource, high-activity signals should be placed near each other and share the large spaces, while low-activity signals share small spaces.

The method is illustrated in Figure 2. There, a bundle contains some signals with high activity (H) and some others with low activity (L). The ordering in Figure 2(b) is superior to Figure 2(a), which is apparently the worst. Wire spacing optimization aiming at minimizing the total power will yield smaller (better) power for configuration 2(b), as compared to 2(a).

Signal ordering is effective when signal activities are known a priori, as in the case of an address bus. In common design, practice signals are laid out

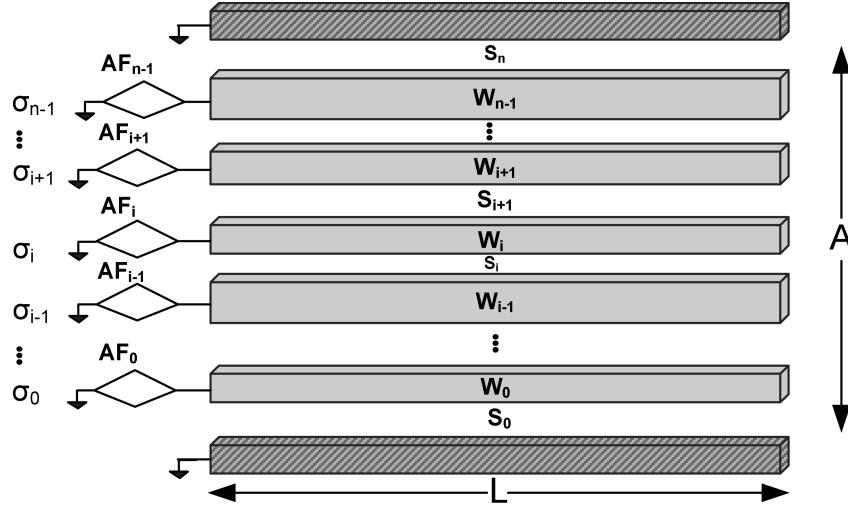


Fig. 1. Model of an interconnect bundle of parallel wires. The  $i$ th wire has activity  $AF_i$ , width  $W_i$ , and spaces to adjacent wires  $S_i$  and  $S_{i+1}$ . The total bundle length is  $L$  and routing area is  $A$ .

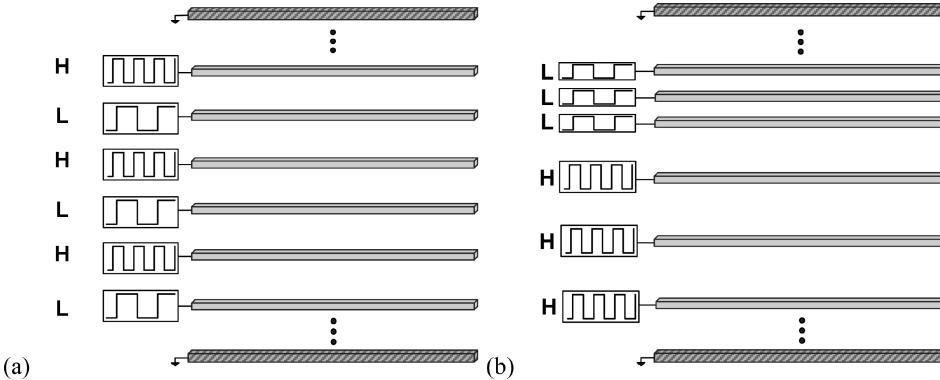


Fig. 2. Space sharing in two interconnect bundle configurations. (a) Interleaved placement of wires with high (H) and low (L) activity, (b) Wires are grouped according to their signal activities.

sequentially from the least significant bit, which is the most active, to the most significant bit, which is typically the least active when sequential addresses are used. However, the ordering that minimizes power consumption is the one where the least significant bits are positioned at the center of the bundle and the two most significant bits are positioned on the two sides of the bus, as illustrated in Figure 3. When some signals in a bundle require shielding, it is beneficial to place them near the sidewalls of the bundle and order the rest of the signals in the central area, according to their activity factors.

Although net-ordering and spacing for delay and cross-talk noise reduction has been discussed widely in the literature [Kahng et al. 1998; Sapatnekar 1996; Cong et al. 2001; Gupta and Kahng 2004; Mui et al. 2004; Moiseev et al. 2006; Jhang et al. 1994], it has been addressed very superficially for power

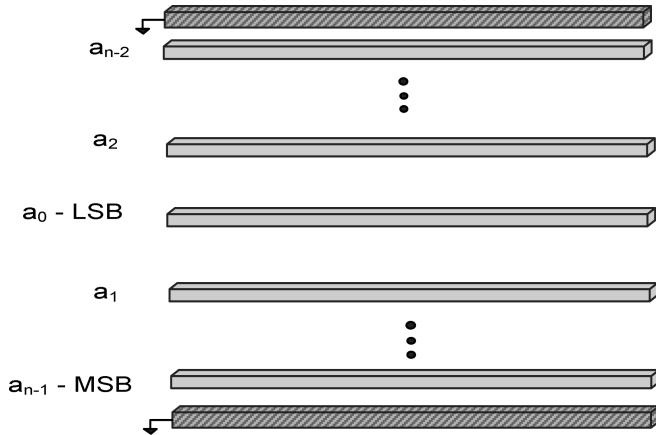


Fig. 3. Optimal ordering of wires in an address bus. The LSB, which has the largest activity, is placed in the middle. The MSB, with lowest activity, is placed near a sidewall (supply voltage wire).

optimization. Coupling capacitance has been addressed explicitly in the context of physical design for minimizing dynamic power in Macii et al. [2003], Arunachalam et al. [2003], Lyuh et al. [2002], Ghoneima and Ismail [2004], and Shin and Sakurai [2001]. Swapping of wires for power reduction was applied in Macii et al. [2003], Lyuh et al. [2002], Shin and Sakurai [2001], and Naroska et al. [2005].

Macii et al. [2003] use a local heuristic approach for inter-wire space allocation, by initially allocating the smallest allowed space and then increasing it until the cross-coupling constraint is satisfied. Although signals are intuitively placed in monotonic order of activity, the heuristic begins with wires with small activity, which would lead to a different result rather than the optimal order shown below.

Lyuh et al. [2002] propose a solution modeled as a bipartite weighted matching problem, implying rescheduling of signals in each clock cycle, which requires additional logic. Moreover, the authors don't solve the space allocation problem.

Shin and Sakurai [2001] also place highly active signals near each other, but their approach has two disadvantages: they divide all signals into several clusters, separated by signals with low activity serving as "shields", and space allocation is not discussed. Such ordering doesn't provide an optimal solution since white-space is not distributed optimally among wires. This happens because highly active signals in one cluster cannot share large spaces with highly-active signals from other clusters.

In the publication by Naroska et al. [2005], the spacing allocation problem for a given signal order is solved optimally. However, the optimal ordering of wires is not used, and a greedy heuristic is employed instead.

In contrast with these papers, we define the optimal signal order as the order that provides the lowest dynamic power dissipation that is achievable by space allocation. A mathematically proven solution, yielding simultaneously the optimal ordering and the optimal space allocation for the signals of a wire bundle, is presented in this article.

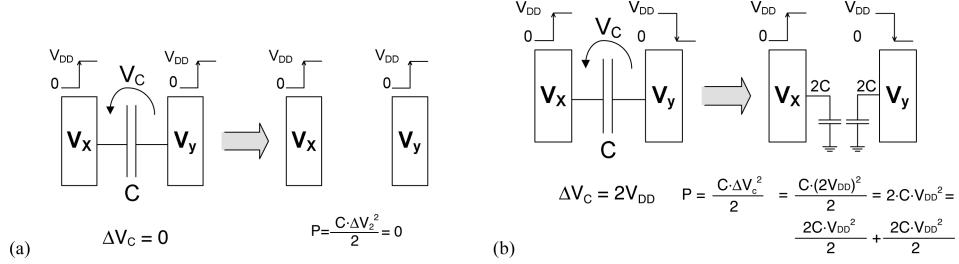


Fig. 4. Miller's theorem for power. (a) Miller's theorem for simultaneous switching of two wires in the same direction—MCF = 0; (b) Miller's theorem for simultaneous switching of two wires in opposite directions—MCF = 2.

## 2. OPTIMAL NET ORDERING AND SPACING FOR POWER MINIMIZATION

Consider a bundle of  $n$  signal nets  $\sigma_0, \dots, \sigma_{n-1}$  residing between two side-walls (wires at fixed locations, connected to  $V_{cc}$  or  $V_{dd}$ ) as shown in Figure 1.  $W_i$ ,  $S_i$  and  $S_{i+1}$ , respectively, denote width and spaces to neighbors of wire  $\sigma_i$ . The length of each wire is  $L$ . The distance  $A$  between the side walls is predefined and needs to satisfy the following constraint:

$$g(\overline{W}, \overline{S}) = \sum_{i=0}^{n-1} W_i + \sum_{i=0}^n S_i = A. \quad (2.1)$$

Assuming full voltage swing at each node, the dynamic power consumed by toggling of wire  $i$  is:

$$P_i = AF_i \left( \alpha L W_i + \delta L \left( \frac{MCF_i^p}{S_i} + \frac{MCF_{i+1}^p}{S_{i+1}} \right) + \gamma L \right) V_{dd}^2 f, \quad (2.2)$$

where  $\alpha, \delta, \gamma$  are coefficients of area, coupling and fringe capacitances,  $V_{dd}$  is supply voltage and  $f$  is the clock frequency. Although fringe, area and coupling capacitances are not fully independent [Stellari and Lacaia 2000], expression (2.2) gives a good first-order approximation for interconnect power.  $AF_i$  is the activity factor of  $\sigma_i$  and  $MCF_i^p$  is the MCF for power (Miller Coupling Factor between the  $i - 1$ th and  $i$ th wire, calculated for power rather than for timing).

The treatment of MCF for power is different than for timing. According to Miller's theorem the simultaneous switching of two adjacent wires in identical or opposite directions yields power MCF of 0 or 2, respectively (see expressions in Figure 4). Assuming that the signals are logically independent, simultaneous transitions in identical and opposite directions are equally likely. Hence, the energy dissipated over multiple simultaneous switching transitions can be calculated using the average MCF for power, which is equal to 1 (same as the MCF for nonsimultaneous switching, when the adjacent signal is stable). Therefore, an average MCF of 1 will be assumed in the derivation below for the sake of simplicity. For the side nets  $\sigma_0$  and  $\sigma_{n-1}$ , the MCF is 1, since the sidewall wires are shields which are always stable.

Rearrangement of (2.2) yields:

$$P_i = AF_i \left( aW_i + b \left( \frac{MCF_i^p}{S_i} + \frac{MCF_{i+1}^p}{S_{i+1}} \right) + P^0 \right), \quad (2.3)$$

where  $a = \alpha LV_{dd}^2 f$ ,  $b = \delta LV_{dd}^2 f$ , and  $P^0 = \gamma LV_{dd}^2 f$ .

The mathematical technique used below to minimize the power is based on a timing-optimization approach described in [Moiseev et al. 2007].

Substituting  $MCF = 1$  in (2.3) and summing over all signals yields the following total interconnect active power:

$$\begin{aligned} P(\overline{W}, \overline{S}) &= \sum_{i=0}^{n-1} P_i = b \sum_{i=1}^{n-1} \frac{AF_{i-1} + AF_i}{S_i} + \frac{AF_0 \cdot b}{S_0} + \frac{AF_{n-1} \cdot b}{S_n} \\ &\quad + a \sum_{i=0}^{n-1} AF_i \cdot W_i + nP^0. \end{aligned} \quad (2.4)$$

Assume for the moment that some order (permutation)  $\pi$  of the signals in the bundle is given. Minimization of (2.4) subject to (2.1) is obtained by differentiating  $P$  and  $g$  by all of their sizing variables. Let's assume that wire widths are allocated in advance according to other design considerations, such as wire delays, electromigration effects or design rules, and therefore they are not part of the optimization. Notice that from a pure power viewpoint, disregarding timing, minimum power is achieved by setting  $W_i = W_{\min}$ ,  $0 \leq i \leq n-1$ . Thus, we differentiate  $P$  and  $g$  only by variables  $S_i$ , which yields the following spacing, minimizing total power for a given order  $\pi$  [Moiseev et al. 2007]:

$$\begin{cases} S_i = \sqrt{\frac{b}{\lambda} (AF_{i-1} + AF_i)}, & 0 < i < n; \\ S_0 = \sqrt{\frac{b}{\lambda} AF_0}; \\ S_n = \sqrt{\frac{b}{\lambda} AF_{n-1}}, \end{cases} \quad (2.5)$$

where  $\lambda$  is a positive constant (Lagrange multiplier).

Let  $\Pi$  denote the set of all wire permutations in the bundle and consider now  $\pi \in \Pi$  as variable. Let  $\pi^*$  denote the permutation for which the optimal wire spacing yields minimum total power among all  $\pi \in \Pi$ . One needs therefore to solve the problem:

Minimize:  $P(\pi, \overline{W}, \overline{S})$ , subject to:  $\sum_{j=0}^{n-1} W_i + \sum_{j=0}^n S_i = A$ .

In this formulation, both signal ordering and wire spacing are optimized simultaneously.

Substitution of (2.5) into (2.4) produces the following expression for the minimal power at a given permutation:

$$\begin{aligned} P &= a \sum_{i=0}^{n-1} AF_i \cdot W_i + \frac{b}{A - \sum_{i=0}^{n-1} W_i} \left( \sum_{i=1}^{n-1} \sqrt{AF_{i-1} + AF_i} + \sqrt{AF_0} + \sqrt{AF_{n-1}} \right)^2 \\ &\quad + nP^0 = P^I + P^{II}, \end{aligned}$$

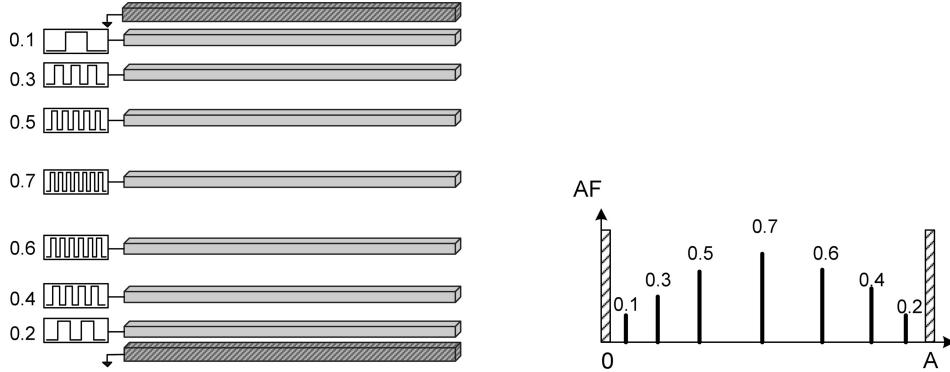


Fig. 5. Symmetric hill ordering by Activity Factors. On the left: an interconnect bundle with wires placed in symmetric hill order and spaced for minimum power: the most active wire is in the middle with maximal spacing and the least active wires are near the walls with minimal spacing. On the right: net activity vs. wire location within the bundle.

where

$$P^I = a \sum_{i=0}^{n-1} W_i \cdot AF_i + nP^0$$

and

$$P^{II} = \frac{b}{A - \sum_{i=0}^{n-1} W_i} \left( \sum_{i=1}^{n-1} \sqrt{AF_{i-1} + AF_i} + \sqrt{AF_0} + \sqrt{AF_{n-1}} \right)^2. \quad (2.6)$$

The term  $P^I$  is invariant for any ordering of the signals. In the term  $P^{II}$ , the indices of adjacent signals interact with each other in the square root expressions, thus making  $P^{II}$  dependent on the order of signals in the bundle. The reason for this interaction is the cross-capacitance between adjacent wires, caused by the space they share with each other. The mathematical properties of the expressions of the kind (2.6) were discussed in the context of minimizing the total sum of weighted delays [Moiseev et al. 2007]. Fortunately, Eq. (2.6) is similar to the order-dependent portion of total weighted delay discussed in Moiseev et al. [2007].

Consequently, the order of wires which minimizes the total bundle power is obtained as follows. Signal nets are sorted in ascending order of activity factors  $AF_0 \leq AF_1 \leq \dots \leq AF_{n-2} \leq AF_{n-1}$ . The sorted set is split into even and odd subsequences  $AF_0 \leq AF_2 \dots$  and  $AF_1 \leq AF_3 \leq \dots$ . By reversing the order of numbers in the odd subsequence it becomes a monotonic decreasing sequence. Finally, the even and the modified (reversed) odd subsequences are concatenated into one sequence. The new sequence thus obtained is said to be ordered in *symmetric hill ordering by activity factor* (as it resembles climbing and descending a symmetric hill). Figure 5 illustrates such an order.

The optimality of symmetric hill order is proven in Moiseev et al. [2007] for the total sum of weighted delays. Since the order dependent portion of the total weighted delay is mathematically similar to (2.6), symmetric hill order is

optimal also for total bundle power minimization. The optimal bundle power can be achieved as follows: first, wires are ordered in symmetric hill order with respect to signal activity factors; then, inter-wire spacing allocation is performed in accordance with Eq. (2.5). The obtained setting yields minimum total dynamic power dissipation in the wire bundle.

### 3. TIMING-AWARE POWER OPTIMIZATION

In a real design flow, the optimization of power without consideration of timing is impractical. In this section, we discuss a timing-aware power optimization technique.

Signal delays are expressed by an Elmore model using simple approximations for wire capacitances, like in (2.2), and wire resistance. The delay of signal  $i$  is given in Wimer et al. [2006] by

$$\begin{aligned} D_i = & \left( R^{dr} + \frac{1}{2} \frac{\beta L}{W_i} \right) \delta L \cdot \left( \frac{MCF_i^d}{S_i} + \frac{MCF_{i+1}^d}{S_{i+1}} \right) + \frac{1}{2} \beta \gamma \frac{L}{W} + \alpha R^{dr} LW_i \\ & + R^{dr} C^L + \frac{1}{2} \alpha \beta L^2 + \frac{\beta L}{W_i} C^L \end{aligned} \quad (3.1)$$

where  $\beta$  is the wire sheet resistance,  $R^{dr}$  is the driver resistance and  $C^L$  is the load capacitance.  $MCF_i^d$  and  $MCF_{i+1}^d$  are Miller Coupling Factors for delay calculations. Although the Elmore model is a first-order approximation and it does not account for input waveform slope [Kahng et al. 1996], it is widely applicable in interconnect optimization due to its high-fidelity property [Boese et al. 1993]. The absolute accuracy of the model can also be improved, by using parameter fitting as described in Abou-Seido et al. [2002]. This model is used in this work because of its simplicity in mathematical analysis, while the delay improvements are verified by SPICE simulations.

Although timing considerations lead to a range of MCF in the interval  $(-1, 3)$  [Chen et al. 2000], usually for timing calculations  $MCF = 2$  is assumed when neighbor wires switch in opposite directions causing increased delays, while  $MCF = 0$  is assumed for same-direction switching and reduced delay. We assume that  $MCF = 1$  for all of the signals, yielding nominal delay values.

Our goal is optimization of power with consideration of timing. The commonly used objective functions incorporating both power and delay are power-delay product or similar multiplicative metrics. However, these functions are not handy for mathematical analysis, and they are not differentiable if maximum delay (worst wire delay) is used. Instead, an objective function based on a weighted sum rather than a product of delay and power is proposed:

$$F = p \sum_{i=0}^{n-1} P_i + q \sum_{i=0}^{n-1} D_i = \sum_{i=0}^{n-1} (pP_i + qD_i), \quad (3.2)$$

where  $p$  and  $q$  are normalization coefficients of power and delay criticality in the optimization. For example, for  $p = 1, q = 0$  pure power optimization will be performed and if  $p = 0, q = 1$ , pure delay optimization is performed. By substituting expressions for  $P_i$  and  $D_i$  and appropriate values for MCF we

Table I. Cross Capacitance Multiplication Coefficients for Different Optimization Problems

Optimization	$\kappa_i$	Meaning of $\kappa_i$
Average delay	$R_i^{dr} + \frac{1}{2} \frac{\beta L}{W_i}$	Effective signal resistance (sum of driver resistance and wire resistance)
Total power	$AF_i \cdot V_{dd}^2 f$	Signal activity factor
Power-delay weighted sum	$pAF_i \cdot V_{dd}^2 f + q \left( R_i^{dr} + \frac{1}{2} \frac{\beta L}{W_i} \right)$	Weighted positive linear combination of activity factor and effective resistance. $p$ and $q$ are weights.

obtain:

$$\begin{aligned} F = & \sum_{i=0}^{n-1} \left( pAF_i \cdot V_{dd}^2 f + q \left( R^{dr} + \frac{1}{2} \frac{\beta L}{W_i} \right) \right) \cdot \delta L \left( \frac{1}{S_i} + \frac{1}{S_{i+1}} \right) \\ & + \sum_{i=0}^{n-1} \left( pAF_i (\alpha LW_i + \gamma L) V_{dd}^2 f + q \left( \frac{1}{2} \beta \gamma \frac{L}{W} + \alpha R^{dr} LW_i + R^{dr} C^L \right. \right. \\ & \left. \left. + \frac{1}{2} \alpha \beta L^2 + \frac{\beta L}{W_i} C^L \right) \right). \end{aligned} \quad (3.3)$$

As can be seen from (3.3), the objective function can be divided to order-dependent (first sum) and order-independent (second sum) parts. The order-dependent part is a sum of cross-capacitances of each net weighted by a linear combination of parameters related to net power (activity factor) and delay (effective resistance of the driver and the wire). Comparison of (3.3) with (2.4) shows that the order-dependent parts of both expressions are similar, apart from coefficients multiplying the cross capacitances:

$$\sum_{i=0}^{n-1} \kappa_i \cdot \delta L \left( \frac{1}{S_i} + \frac{1}{S_{i+1}} \right),$$

where the appropriate coefficients are denoted by  $\kappa_i$ . The analysis in Moiseev et al. [2007] and in Section 2 show that the optimal order of signals actually depends on these coefficients. Table I summarizes the different optimization cases discussed in Moiseev et al. [2007] and in this article.

In all cases, the optimal order is Symmetric Hill in accordance with the corresponding  $\kappa_i$ . It is readily seen from the table that the optimal order for power-delay sum optimization will be similar to that of total power optimization or average delay optimization, if monotony of the corresponding  $\kappa_i$  is preserved. This can be summarized in following theorem:

**THEOREM 1.** *If any two signals  $\sigma_i$  and  $\sigma_j$  satisfy*

$$AF_i \geq AF_j \Leftrightarrow \left( R_i^{dr} + \frac{1}{2} \frac{\beta L}{W_i} \right) \geq \left( R_j^{dr} + \frac{1}{2} \frac{\beta L}{W_j} \right), \quad (3.4)$$

*then symmetric hill order by activity factors is optimal with respect to effective signal resistances and vice-versa. Moreover, it is optimal with respect to any linear combination of the two.*

**PROOF.** The linear combination statement follows directly from the optimality for both power and delay as follows: if any two signals  $\sigma_i$  and  $\sigma_j$  simultaneously satisfy  $AF_i \geq AF_j$  and  $(R_i^{dr} + \frac{1}{2} \frac{\beta L}{W_i}) \geq (R_j^{dr} + \frac{1}{2} \frac{\beta L}{W_j})$ , then  $\kappa_i = pAF_i + q(R_i^{dr} + \frac{1}{2} \frac{\beta L}{W_i}) \geq pAF_j + q(R_j^{dr} + \frac{1}{2} \frac{\beta L}{W_j}) = \kappa_j$ .  $\square$

Fortunately, the relation (3.4) is satisfied in most practical cases. Note that reducing the driver size and wire width of highly active signals helps to reduce interconnect power.

In practice, power optimization described above can be applied in the early design stages, when signal timing criticalities are not known yet. In later design stages, power optimization with consideration of signal criticality is of interest. Power-delay optimization that considers power-delay and timing criticality can be captured by introducing a criticality factor for each signal, calculated as follows:

$$\kappa_i = \chi_i AF_i \cdot V_{dd}^2 f, \quad (3.5)$$

where  $\chi_i$  is a normalized measure of timing criticality of the  $i$ th signal. The least critical signal will have  $\chi_i = 1$  while all the others will have  $\chi_i > 1$ . Substituting (3.5) in (2.5) and (2.6) yields

$$\begin{cases} S_i = \sqrt{\frac{b}{\lambda} (\chi_{i-1} AF_{i-1} + \chi_i AF_i)}, & 0 < i < n; \\ S_0 = \sqrt{\frac{b}{\lambda} \chi_0 AF_0}; \\ S_n = \sqrt{\frac{b}{\lambda} \chi_{n-1} AF_{n-1}} \end{cases} \quad (3.6)$$

$$P^{II} = \frac{b}{A - \sum_{i=0}^{n-1} W_i} \left( \sum_{i=1}^{n-1} \sqrt{\chi_{i-1} AF_{i-1} + \chi_i AF_i} + \sqrt{\chi_0 AF_0} + \sqrt{\chi_{n-1} AF_{n-1}} \right)^2. \quad (3.7)$$

Thus, since the optimal order now is symmetric hill with respect to  $\kappa_i = \chi_i AF_i \cdot V_{dd}^2 f$ , timing-critical and highly active signals will be placed close to each other and will share large spaces thus reducing the shared cross-capacitance load.

#### 4. EXPERIMENTAL RESULTS

The following experiments demonstrate the optimization techniques described above, applied to 65-nm process technology.

*Experiment 1.* Impact of total area allocated for the wire bundle on power reduction by reordering.

The routing pitch of a given layer  $X = W_{\min} + S_{\min}$  is defined as the sum of minimal width and minimal space (usually they are equal). When the area allocated to an  $n$ -signal bundle is  $A = nW_{\min} + (n+1)S_{\min}$ , wire reordering will not reduce power since the wire to wire spacing must be always minimal, regardless of their order in the bundle. On the other hand, allocating excessive bundle width that allows very large spacing between any two adjacent wires makes the

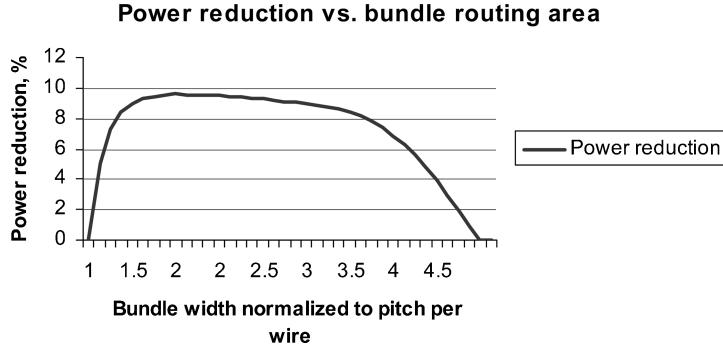


Fig. 6. Power reduction vs. bundle width. For width less than a single pitch per wire or larger than 4.7 pitches per wire (on average) the power reduction is zero in this example, since all spaces must be identical in these ranges because of maximum or minimum spacing constraints.

power almost insensitive to wire ordering. Setting bundle width between these two extreme cases enables significant power reduction, as demonstrated in the following experiment, which uses the following parameters: bundle length  $L = 300 \mu\text{m}$ , 4th level metal layer with  $W_{\min} = S_{\min} = 0.14 \mu\text{m}$  and  $n = 5$ . Different allocations of bundle width were made covering the range from  $11W_{\min}$  up to  $50W_{\min}$  (i.e., from 1 pitch per wire, up to 5 pitches per wire). Widths of all wires were kept minimal during the experiment. For every value of bundle area, 100 random sets of 5 activity factors were drawn and for each set the optimal ordering and spacing were performed. Then average power reduction in comparison with the initial (random) configuration was calculated. The results are shown in Figure 6. There, the maximum reduction is 10%, achieved at allocation of 2 pitch/wire. Since wire width and spacing of twice the minimum design rule is a very common setting in VLSI interconnect bus layouts, ordering and spacing for power reduction is practically useful in this technology.

*Experiment 2.* Power optimization by net reordering and spacing in an industrial design.

This experiment was conducted on a circuit block from a high-end microprocessor designed in 65nm process technology. Activity factors of signals were derived using industrial tools that run a suite of benchmark test cases on a representative block of the design. Figure 7 shows that 95% of the signals have activity of less than 0.2. A typical layout snapshot is shown in Figure 8, exhibiting typical interconnect signal bundles extending across the block. Some bundles are shown at a larger zoom. Dark-colored signals have low activity while light ones are of higher activity.

The specific parameters of each magnified bundle are shown in Table II.

For each of these bundles, ordering and spacing optimization have been performed. First, spacing optimization for the original signal order was performed and the power reduction was recorded. Then, the bundles were reordered as a symmetric hill according to the underlying activity factors and spacing optimization was re-invoked, and the power reduction was recorded again. Wire widths were not changed in both optimizations. As expected, the second space

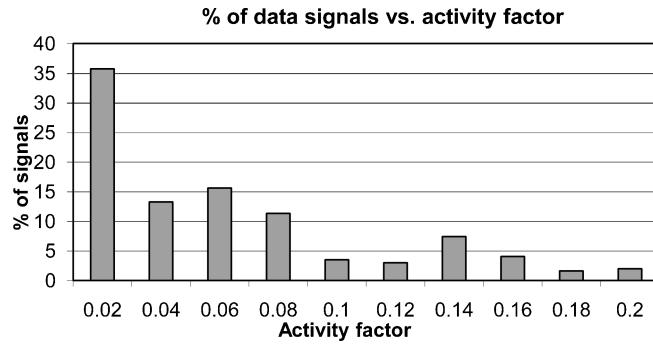


Fig. 7. Distribution of data signals by activity factor for signals with activity less than 0.2.

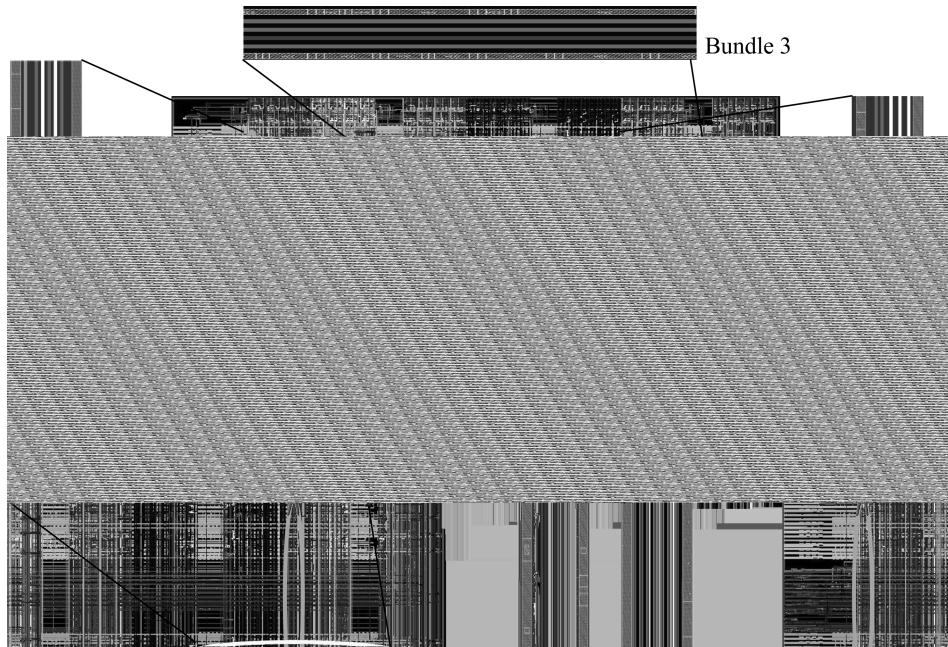


Fig. 8. Snapshot of a processor data path block. Signals are shaded according to their activity factors from 0 (black) to 0.2 (white). Bundles chosen for optimization are shown on the picture.

optimization (after reordering) yielded larger reduction than the first (which used the original signal order). This improvement in power reduction is attributed to wire ordering. Results are presented in Figure 9. In the chart, the bottom part of each bar is the power reduction as a result of the first spacing optimization, and the top part of each bar shows the additional power saving attributed to the optimal reordering of the wires. All savings are calculated with regard to the total power of the bundle in its original spacing and ordering as drawn in the processor layout. The total optimization impact varies from 9% to 37% with 17% on average. The average gain attributed to spacing only is 12%, while average gain attributed to ordering is 5%. The difference in power

Table II. Parameters of Bundles Derived from the Layout

Number of bundle (number of signals in brackets)	1 (6)	2 (6)	3 (4)	4 (5)	5 (5)
Metal layer	3	3	4	2	2
Bundle length, $\mu\text{m}$	150	184	173	96	98
Bundle width, $\mu\text{m}$	1.77	2.105	2.94	1.7	1.7
Signal activity factors	0.064; 0.014; 0.023; 0.097; 0.005; 0.014	0.066; 0.063; 0.062; 0.065; 0.178; 0.204	0.025; 0.045;	0.059; 0.205; 0.073; 0.159; 0.066	0.158; 0.06; 0.066; 0.075; 0.204

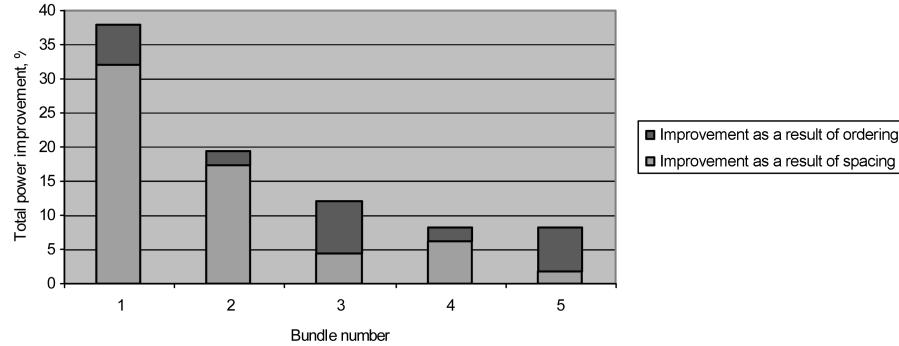
**Bundle power improvement as result of spacing and ordering optimization**

Fig. 9. Power reduction as a result of spacing and ordering optimizations.

savings for different bundles can arise from various reasons: bundle density, initial bundle spacing, range of bundle signal activity factors, and closeness of the initial signal order to the optimal one.

*Experiment 3.* Power optimization by net reordering and spacing with timing consideration.

This experiment demonstrates how consideration of signal timing criticality affects the optimal order of signals. Consider a  $3.3\text{-}\mu\text{m}$  width and  $300\text{-}\mu\text{m}$  length bundle of 6 signals. All drivers have  $400\Omega$  output resistance, and all load capacitances are  $10\text{ fF}$ . The bundle consists of different signals whose characteristics are summarized in Table III below. Signal 1 is timing-critical with required time of  $20\text{ psec}$ , while signals 2 and 4 are noncritical with required time of  $40\text{ psec}$ . Signals 1, 2, and 4 have small activity factors. Signals 3, 5, and 6 are outputs of a fast 3-bit counter with a required time of  $25\text{ psec}$  and high activity factors. The cross section in Figure 10(a) shows the initial setting of the bundle, where all signals are equally spaced. Signal 6, which has the largest activity factor of 0.5, is allocated with the same spaces as signal 1 with the smallest activity factor of 0.05.

The outcome of the symmetric hill ordering of the bundle in accordance with activity factors, followed by spacing for power minimization is shown in Figure 10(b). As expected, signal 6 is positioned at the center and is allocated

Table III. Characteristic of Bundle Wires in Experiment 3

Net Number	Net Purpose	Required Time
1	highly critical signal	20 psec
2,4	non-critical signals	40 psec
3,5,6	fast 3-bit counter	25 psec

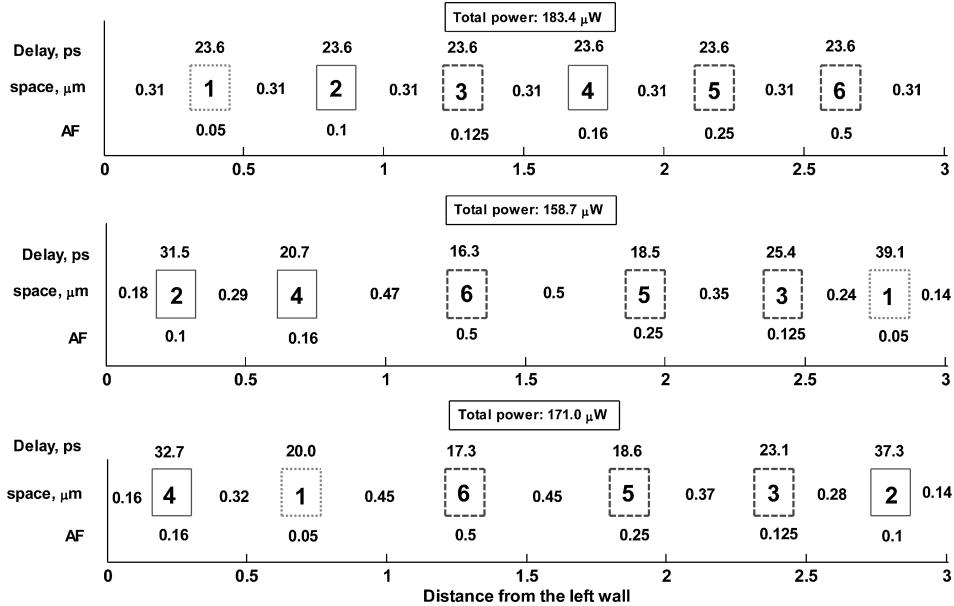


Fig. 10. Cross section of a bundle: (a) Initial layout—all wires are uniformly spaced; (b) The wires are placed in symmetric hill order of activity factors and spaced for minimum power. The critical wire #1 (dotted) is near the wall and its delay is increased by 40%; (c) Timing aware ordering and spacing optimization: critical wire #1 is placed near the most active wire and its delay decreased by about 50%.

with the largest spaces. Total bundle power is reduced by 13.5% compared to the initial setting, but the delay of the signals 1 is increased by 40%, which severely violates its required time. This happened because the optimization was aware only to its smallest activity factor but not to its timing criticality.

Finally, the power-delay combined ordering is performed. Criticality factors are assigned to the signals reflecting their required time. The less critical factors are  $\kappa_2 = \kappa_4 = 1$ . The signals of the fast counter that are more critical are weighted by  $\kappa_3 = \kappa_5 = \kappa_6 > 1$ . For the most critical signal, we assign  $\kappa_1 \gg 1$ . The outcome of the timing-criticality aware symmetric hill order followed by spacing optimization is shown in Figure 10(c). There, signal 1 got closer to the center with higher spaces as compared with Figure 10(b), and its required time was met. Signal 3 (which didn't meet its required time in Figure 10(b)) is also improved. All the other signals meet their required time specifications. As expected, this timing improvement came at the expense of some power reduction, which is now 6.8% compared with the initial setting. The results are

Table IV. Experiment 3—Summary of Results

	Net Slacks, [ps]						Total Bundle Power, [ $\mu\text{W}$ ]	Power in % of Initial
	1	2	3	4	5	6		
Initial placement	-3.6	+16.4	+2.4	+16.4	+2.4	+2.4	183.4	100%
Ordering and spacing for power	-19.1	+8.5	-0.4	+19.3	+6.5	+8.7	158.7	86.5%
Criticality-aware ordering and spacing	0	+2.7	+1.9	+7.3	+6.4	+7.7	171.0	93.2%

Table V. Increase of Metal and Via Aspect Ratios in Future Technologies

Year of production	2005	2007	2009	2011	2013	2017
Metal 1½ pitch, nm	90	68	52	40	32	20
Metal AR ( $\text{AR}_m$ )	1.7	1.8	1.8	1.8	1.9	2
Via AR ( $\text{AR}_v$ )	1.5	1.6	1.6	1.6	1.7	1.8
$P^H/P^I$	0.111	0.118	0.118	0.118	0.126	0.133

summarized in Table IV, demonstrating how the criticality coefficient method takes into account both net activity and timing criticality.

## 5. DEPENDENCE OF OPTIMIZATION IMPACT ON PROCESS TECHNOLOGY

As lateral feature sizes decrease with technology advancement, an important question is how bundle power will be affected by net ordering in future manufacturing process technology generations. Let's analyze the ratio  $P^H/P^I$ . The larger this ratio is, the more effective wire ordering is. Substituting expressions for  $a$ ,  $b$  and  $P^0$  in (2.3) into (2.6) yields:

$$\frac{P^H}{P^I} = \frac{\left( \sum_{i=1}^{n-1} \sqrt{AF_{i-1} + AF_i} + \sqrt{AF_0} + \sqrt{AF_{n-1}} \right)^2}{A - \sum_{i=0}^{n-1} W_i} \cdot \frac{\delta L}{\alpha L \cdot \sum_{i=0}^{n-1} AF_i W_i + \gamma n L}. \quad (6.1)$$

The term  $k = (\sum_{i=1}^{n-1} \sqrt{AF_{i-1} + AF_i} + \sqrt{AF_0} + \sqrt{AF_{n-1}})^2$  is independent of layout and technology, hence

$$\frac{P^H}{P^I} = \frac{k}{A - \sum_{i=0}^{n-1} W_i} \cdot \frac{\delta}{\alpha \cdot \sum_{i=0}^{n-1} AF_i W_i + \gamma n} \quad (6.2)$$

Denote by  $\overline{AF} = \frac{1}{n} \sum_{i=0}^{n-1} AF_i$  the average activity factor, and assume for convenience that all wires have the same width  $W$ . Substitution into (6.2) yields

$$\frac{P^H}{P^I} = \frac{k}{n(A - nW)} \cdot \frac{\delta}{\alpha \overline{WF} + \gamma}. \quad (6.3)$$

Wire ordering becomes more effective as the ratio (6.3) becomes larger. Let's check dependence of (6.3) on technology parameters.

For a first-order analysis, the ratios  $\alpha/\delta$  and  $\gamma/\delta$  are approximated by  $1/t_{ox} H$  and  $1.06/\sqrt{t_{ox} H}$  [ITRS 2005], where  $H$  is metal line height and  $t_{ox}$  is height of inter-layer dielectric. Assuming a scaling factor of  $s$  for the bundle width and

the wire width, the ratio can be expressed as:

$$\begin{aligned} \frac{P^H}{P^I} &= \frac{k}{A - nW} \cdot \left( \frac{1}{t_{ox}H} W \cdot \overline{AF} + \frac{1.06n}{\sqrt{t_{ox}H}} \right)^{-1} \\ &= \frac{k}{\frac{A}{W} - n} \cdot \left( \frac{1}{AR_m \cdot AR_v} \cdot \overline{AF} + 1.06n \cdot \frac{1}{\sqrt{AR_m \cdot AR_v}} \right)^{-1}. \end{aligned} \quad (6.4)$$

where  $AR_m = \frac{H}{W_s}$  and  $AR_v = \frac{t_{ox}}{W_s}$  are aspect ratios (thickness of material/minimum width) of a metal line and a via, respectively.

Table V has been derived from ITRS reports [2005], predicting these values for several technology generations ahead. Both are increasing from generation to generation. The parameters of Eq. (6.4) such as  $A$ ,  $W$ ,  $n$  and  $\overline{AF}$  were derived from design data shown in Table I. Substitution into (6.4) shows that the ratio  $P^H/P^I$  steadily increases, thus making the ordering optimization more effective with process technology evolution (last row of Table V).

## 6. CONCLUSION

In this article, we showed how the total switching power of interconnect wire bundles can be reduced while taking into account the timing criticality of its signals. The power was reduced by simultaneous net spacing and optimally-proven ordering in accordance with signal activity factors. The optimal order of the signals within the bundle depends only on their activity factors and is of symmetric hill form. Numerical experiments have shown that the effectiveness of wire reordering strongly depends on the total width allocated for the wire bundle. The largest reduction was achieved for bundles with an average width of two metal pitches per signal. The power saving achieved by the spacing and ordering combined optimization that has been performed on industrial layouts of 65-nm process technology ranged from 9% to 37%. Although in terms of the entire power consumption this turns into a smaller percentage, it is still significant. It is highly recommended to apply wire ordering optimization at the early stages of design and employ it as a guideline for the routing tool in use.

## REFERENCES

- ABOU-SEIDO, A. I., NOWAK, B., AND CHU, C. 2002. Fitted elmore delay: A simple and accurate interconnect delay model. In *Proceedings of the IEEE International Conference on Computer Design*, IEEE Computer Society Press, Los Alamitos, CA, 422–427.
- ARUNACHALAM, R., ACAR, E., AND NASSIF, S. R. 2003. Optimal shielding/spacing metrics for low power design. In *Proceedings of IEEE Computer Society Annual Symposium on VLSI*, IEEE Computer Society Press, Los Alamitos, CA, 167–172.
- BARKE, E. 1988. Line-to-ground capacitance calculation for VLSI—A comparison. *IEEE Trans. Computer-Aided Des. VLSI*, 7, 2, 295–298.
- BENINI, L., DE MICHELI, G., MACHI, E., SCIUTO, D., AND SILVANO C. 1998. Address bus encoding techniques for system-level power optimization. In *Proceedings of DATE 1998*, 861–866.
- BERTOZZI, D., BENINI, L., AND DE MICHELI, G. 2002. Low power error resilient encoding for on-chip data buses. In *Proceedings of DATE 2002*, 102–109.
- BOESE, K. D., KAHNG, A. B., MCCOY, B. A., AND ROBINS, G. 1993. Fidelity and near-optimality of elmore-based routing constructions. *Dig. Technical Paper ICCAD*, pp. 81–84.
- CHA, M., LYUH, C., AND KIM, K. 2006. Low power bus encoding with crosstalk delay elimination. *IEE Proceedings: Computers and Digital Techniques* 153, 2, 93–100.

- CHEN, P., KIRKPATRICK, D. A., AND KEUTZER, K. 2000. Miller factor for gate-level coupling delay calculation. In *Proceedings of ICCAD 2000*, 68–74.
- CONG, J., HE, L., KOH, C. K., AND PAN, Z. 2001. Interconnect sizing and spacing with consideration of coupling capacitance. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 20, 9, 1164–1169.
- GHONEIMA, M. AND ISMAIL, Y. 2004. Delayed line bus scheme: a low-power bus scheme for coupled on-chip buses. In *Proceedings of 2004 International Symposium on Low Power Electronics and Design*, 66–69.
- GUPTA, P. AND KAHNG, A. 2004. Wire swizzling to reduce delay uncertainty due to capacitive coupling. In *Proceedings of IEEE International Conference on VLSI Design*, 431–436.
- HSIEH, C. AND PEDRAM, M. 2000. Architectural power optimization by bus splitting. In *Proceedings of DATE 2000*, 612–616.
- ITRS REPORT, 2005. Available online <http://www.itrs.net/reports.html>
- JHANG, K., HA, S., AND JOHN, C. 1994. A segment rearrangement approach to channel routing under the crosstalk constraints. In *Proceedings of the Asia-Pacific Conference on Circuits and Systems*, 536–541.
- KAHNG, A., MASUKO, K., AND MUDDU, S. 1996. Analytical delay models for VLSI interconnects under ramp input. In *Proceedings of ICCAD 1996*, 30–36.
- KAHNG, A., MUDDU, S., SARTO, E., AND SHARMA, R. 1998. Interconnect tuning strategies for high-performance ICs. In *Proceedings of DATE 1998*, 471–478.
- KIM, K., BAEK, K., SHANBHAK, N., LIU, C., AND KANG, S. 2000. Coupling-driven signal encoding scheme for low-power interface design. In *Proceedings of ICCAD 2000*, 318–321.
- LYUH, C., KIM, T., AND KIM, K. 2002. Coupling-aware high-level interconnect synthesis for low power. In *Proceedings of ICCAD 2002*, 609–613.
- MACHI, E., PONCINO, M., AND SALERNO, S. 2003. Combining wire swapping and spacing for low-power deep-submicron buses. In *Proceedings of the 13th ACM Great Lakes symposium on VLSI*, ACM, New York, 198–202.
- MAGEN, N., KOLODNY, A., WEISER, U., AND SHAMIR, N. 2004. Interconnect-power dissipation in a microprocessor. In *Proceedings of the 2004 International Workshop on System Level Interconnect Prediction*, 7–13.
- MOISEEV, K., WIMER, S., AND KOLODNY, A. 2006. Timing optimization of interconnect by simultaneous net-ordering, wire sizing and spacing. In *Proceedings of IEEE International Symposium on Circuits and Systems*, IEEE Computer Society Press, Los Alamitos, CA, 329–332.
- MOISEEV, K., WIMER, S., AND KOLODNY, A. 2007. On optimal ordering of signals in parallel wire bundles. *Integration—the VLSI journal*, Vol. 41, Issue 2, 253–268.
- MUI, M. L., BENERJEE, K., AND MEHORTRA, A. 2004. A global interconnect optimization scheme for nanometer scale VLSI with implications for latency, bandwidth, and power dissipation. *IEEE Trans. Electr. Dev.* 51, 2, 195–203.
- NAROSKA, E., RUAN, S.-J., AND SCHWIEGELSHOHN, U. 2005. An efficient algorithm for simultaneous wire permutation, inversion, and spacing. In *Proceedings of International Symposium on Circuits and Systems*, 109–112.
- SAPATNEKAR, S. S. 1996. Wire sizing as a convex optimization problem: Exploring the area delay tradeoff. *IEEE Trans. Comput.-Aided Design of Integrated Circuits and Systems* 15, 8, 1001–1011.
- SHANG, L., PEH, L., AND JHA, H. 2002. Power-efficient interconnect networks: Dynamic voltage scaling with links. *Comput. Architect. Lect.* 1, 1, 6–10.
- SHIN, Y. AND SAKURAI, T. 2001. Coupling-driven bus design for low-power application-specific systems. In *Proceedings of the 38th Conference on Design Automation*, 750–753.
- SELLARI, F. AND LACAITA, A. L. 2000. New formulas of interconnect capacitances based on results of conformal mapping method. *IEEE Trans. Electron Dev.*, 222–231.
- WIMER, S., MICHAELY, S., MOISEEV, K., AND KOLODNY, A. 2006. Optimal bus sizing in migration of processor design. *IEEE Trans. Circ. Syst. – I*, 53, 5, 1089–1100.
- ZANG, H., GEORGE, V., AND RABAAY, J. M. 2000. Low-swing on-chip signaling techniques: Effectiveness and robustness. *IEEE Trans. VLSI Syst.* 8, 3, 264–272.

Received June 2007; revised June 2008; accepted June 2008