Using well-solvable minimum cost exact covering for VLSI clock energy minimization

Shmuel Wimer a,b,*, Doron Gluzer a, Uri Wimer c

a Engineering Faculty, Bar-Ilan University, Ramat-Gan 52900, Israel
b EE Dept., Technion – Israel Institute of Technology, Haifa 32000, Israel
c Mathematics Faculty, Hebrew University of Jerusalem, Jerusalem, Israel

A R T I C L E   I N F O

Article history:
Received 8 October 2013
Received in revised form 14 February 2014
Accepted 31 May 2014
Available online 10 June 2014

Keywords:
Well-solvable problems
Exact covering
Perfect graph matching
VLSI clocking
Multi-bit flip-flop

A B S T R A C T

To save energy of VLSI systems flip-flops (FFs) are grouped in Multi-Bit Flip-Flop (MBFF), sharing a common clock driver. The energy savings strongly depend the grouping. For 2-bit MBFFs the optimal grouping turns into a minimum cost perfect graph matching problem. For k-bit MBFFs the optimal grouping turns into a minimum cost exact k-covering problem. We show that due to their special setting that is based on the FFs’ data toggling probabilities, those problems are well-solvable in $O(n \log n)$ time complexity.

© 2014 Elsevier B.V. All rights reserved.

1. VLSI energy savings by multi-bit flip-flop grouping

One of the major energy consumers in computing, communication and consumer electronics and other devices is the system’s clock signal, typically responsible for 30%–70% of the total switching energy [13]. Flip-flops (FFs) are the heart of digital systems, used to synchronize their operation and store the system’s state. To drive the FFs, a clock signal is distributed across the chip through a clocking network. FFs consume most of the clock energy. Within a FF, most of the energy is consumed by its internal clock driver. For simplicity, non-essential VLSI design details are ignored, and the interested reader can find those in any VLSI design textbook (e.g. [15]).

k-bit data is usually stored in k individual FFs, where each of those has its own internal clock drivers. In an attempt to reduce the clock energy, a technique called Multi-Bit Flip-Flop (MBFF) has lately been adopted by the VLSI industry [10,4]. A k-bit MBFF combines several FFs integrated in a single entity, such that a common clock driver is used for all the k internal FFs rather than k drivers. The energy savings achieved by using MBFFs is considerable, and may reach up to 20% of the entire system’s energy. The savings depend on the average (expected) data toggling probability $p$ of the individual FFs, called data toggling probability, switching probability, or shortly probability. We use those terms interchangeably. By definition, there is $0 \leq p \leq 1$, where $p = 0$ when the data is never toggling and $p = 1$ when the data is toggling at every clock cycle.

Fig. 1 shows the energy ratio of two and four individual FFs to that of 2-bit and 4-bit MBFFs, respectively. To find the energy savings, we divide the energy difference between $k$ individual FFs and $k$-bit MBFFs, by the energy of the $k$ individual FFs. For small $p$ it shows savings of $(1.6 - 1)/1.6 = 35\%$ for $k = 2$ and $(2.2 - 1)/2.2 = 55\%$ for $k = 4$. For high $p$ the savings is $(1.18 - 1)/1.18 = 15\%$ for $k = 2$ and $(1.3 - 1)/1.3 = 23\%$ for $k = 4$. In typical VLSI systems $0 < p < 0.2$, so high savings is expected.

Combining MBFFs with Data-Driven Clock Gating (DDCG) considerably increases its energy savings. Ordinarily, FFs receive the clock signal regardless of whether or not their data will toggle in the next cycle. In DDCG the clock signal driving a FF is disabled (gated) when the FF’s state is not subject to change (toggleswitch) in the next clock cycle [7,17]. Due to the high hardware overhead involved in generating those signals, it was suggested to group several FFs and derive a joint disabling signal for those. The group size $k$ yielding minimum energy depends on the toggling probabilities [17]. The problem of what FF should belong to what group so that the energy is minimized was studied in [18]. It was shown that under energy model based on the 0/1 toggling
correlation of the FFs, the problem is NP-hard, and a practical heuristic solution based on Minimum Cost Perfect Graph Matching (MCPM) was devised [16].

Applying DDCG in MBFF design methodology was proposed in [5]. However, the grouping in [5] and in other MBFF works [11,19,14] was not aware of the data toggling probabilities and correlations, thus a big amount of potential energy savings was left untreated. The work in [16] used toggling correlation to derive the optimal FFs grouping for DDCG. It required huge data of 0/1 toggling vectors of all the FFs, obtained by simulations, which is a serious design burden. Furthermore, the corresponding optimization problem is NP-hard as mentioned before, and heuristic solution was thus proposed.

In this paper we simplify the optimal grouping formulation by considering FF probabilities rather than their 0/1 toggling vectors. The simplification implies an optimization that is a kind of minimal cost exact k-covering problem, where for \( k = 2 \) it turns into MCPM, formulated as follows. Given \( n \) real numbers (data toggling probabilities of FFs) \( p_i \in [0, 1], n \) even, \( 1 \leq i \leq n \), \( p_1 \leq p_2 \leq \cdots \leq p_n \), find a perfect matching (\( S, t \)), \( 1 \leq j \leq n/2 \), of the integers \( 1, 2, \ldots, n \), minimizing the following energy loss expression (discussed in Section 3)

\[
\sum_{j=1}^{n/2} p_j \left( 1 - p_j \right) + p_j \left( 1 - p_j \right).
\]

For \( k > 2 \), let \( n \) be an integer multiple of \( k \). Consider the partitioning (called also exact \( k \)-covering) of \( \{1, 2, \ldots, n\} \) into \( n/k \) disjoint subsets \( S^i \subset \{1, 2, \ldots, n\}, |S^i| = k, \bigcup_{i=1}^{S^i} = \{1, 2, \ldots, n\} \). Denote by \( (A_m, B_m) \) the partitioning of \( S^i \), \( 1 \leq m \leq \binom{n}{k} \) such that \( A_m = \{1, 2, \ldots, r\}, B_m = \{r+1, \ldots, n\}, A_m \cap B_m = \emptyset \). The minimum cost exact \( k \)-covering problem (discussed in Section 4) is to find those \( S^i \), \( 1 \leq j \leq n/k \), minimizing the energy loss expression

\[
\sum_{i=1}^{k} \sum_{m=1}^{r} \sum_{j \in A_m} p_j \prod_{j \in B_m} (1 - p_j).
\]

While the minimum cost exact covering problem is NP-hard in general [9], we prove that in our special setting it is well-solvable, requires only sorting. Well-solvable cases of hard combinatorial optimization problems are well-known and have been studied by many works. Just a few to mention are special cases of Traveling Salesman Problems (TSP) [2], Steiner tree problems [1], Quadratic Assignment Problems (QAP) [3] and Bounded Knapsack Problem (BKP) [6]. Well-solvable quadratic assignment problems was used for VLSI interconnect design optimization [8].

The rest of the paper is organized as follows. Section 2 presents the MBFF energy consumption model. Section 3 shows that the 2-bit MBFF optimal pairing is MCPM problem, well-solvable by sorting of FFs’ toggling probabilities. Section 4 shows that the \( k \)-bit MBFF optimal grouping is a minimum cost exact \( k \)-covering problem, that it is also well-solvable by sorting of FFs’ toggling probabilities.

2. Energy consumption of multi-bit flip-flops

The energy \( E_1 \) consumed by an ordinary 1-bit FF grows with its toggling probability \( p \) as follows:

\[
E_1(p) = \alpha_1 + \beta_1 p.
\]

The parameter \( \alpha_1 \) is the energy of the FF’s internal clock driver, and the parameter \( \beta_1 \) is the energy of data toggling. For 2-bit MBFF there are three possible scenarios: none of the FFs toggle, a single FF toggles and both FFs toggle. Assuming toggling independency, the energy consumption \( E_2 \) is

\[
E_2(p) = \alpha_1 (1 - p)^2 + 2(\alpha_2 + \beta_2) p (1 - p) + (\alpha_2 + 2\beta_2) p^2
\]

\[\equiv \alpha_2 + 2\beta_2 p.\]

The parameter \( \alpha_2 \) is the energy of the internal clock driver which drives the two FFs, and the parameter \( \beta_2 \) is of data toggling energy of one bit in the 2-bit MBFF. The energy savings factor \( 2E_1(p) / E_2(p) \) is shown in Fig. 1. Obviously, the lower the data toggling probability is, the higher the savings factor is.

For the general case of \( k \)-bit MBFF, let \( \alpha_k \) be the energy of the MBFF’s internal clock driver driving its \( k \) FFs, and let the parameter \( \beta_k \) be the data toggling energy of one bit in the \( k \)-bit MBFF. Considering all the combinations of toggling FFs, the energy consumption \( E_k \) is

\[
E_k(p) = \sum_{j=0}^{k} \binom{k}{j} p^j (1 - p)^{k-j}.
\]

Rearrangement of (3) yields

\[
E_k(p) = \alpha_k \sum_{j=0}^{k} \binom{k}{j} p^j (1 - p)^{k-j} + \beta_k \sum_{j=0}^{k} \binom{k}{j} j p^j (1 - p)^{k-j}.
\]

The equality \( \sum_{j=0}^{k} \binom{k}{j} j p^j (1 - p)^{k-j} = kp \) in (4) follows from \( j \binom{k}{j} = k \binom{k-1}{j-1} \). The energy savings factor \( 4E_1(p) / E_4(p) \) is shown in Fig. 1.

3. Optimal FF grouping of 2-bit MBFF

Let \( FF_i \) and \( FF_j \) toggle independently of each other with probabilities \( p_i \) and \( p_j \), respectively. We denote by \( FF_{i,j} \) their grouping (pairing) in the formation of a 2-bit MBFF. Similar to (2), the energy \( E_{i,j} \) consumed by \( FF_{i,j} \) is \( E_{i,j} = \alpha_2 + \beta_2 (p_i + p_j) \). For \( FF_i, FF_j, FF_k \) and \( FF_{i,j} \), paired in two MBFFs \( FF_{i,j} \) and \( FF_{k,l} \), the energy consumption is \( E_{i,j} + E_{k,l} = 2\alpha_2 + \beta_2 (p_i + p_j + p_k + p_l) \), which is independent of the pairing.

Pairing considerably affects the energy consumption when DDCG is applied. Recall that with DDCG the clock pulse is disabled when the data of a FF will not change (toggling) in the next clock cycle. Since in MBFF the clock signal is common to all FFs, when none of \( FF_i \) and \( FF_j \) is toggling, the clock pulse of \( FF_{i,j} \) is disabled and its clock driver does not waste any energy. When both \( FF_i \) and \( FF_j \) are toggling, the clock pulse of \( FF_{i,j} \) is enabled and the energy of the clock driver is fully useful, hence no waste occurs. Energy waste occurs when one FF is toggling, while its counterpart does.
not. There, the common clock pulse is enabled and is driving both FFs, whereas only one needs it, thus causing a waste $W(i,j)$ of half of the clock driver energy,

$$W(i,j) = \frac{\alpha_2}{2} \left[ p_j (1 - p_j) + p_i (1 - p_i) \right] = \frac{\alpha_2}{2} \left( p_i + p_j - 2p_ip_j \right). \tag{5}$$

We are interested in the minimization of $W(i,j)$. Applying DDCG to FF$_{i,j}$ and FF$_{k,b}$, the following energy waste results in

$$W(i,j) + W(k,b) = \frac{\alpha_2}{2} \left[ p_j (1 - p_j) + p_i (1 - p_i) \right] + p_k (1 - p_k) + p_i (1 - p_i) = \frac{\alpha_2}{2} \left[ p_j + p_k + p_i - 1 \right] + \frac{\alpha_2}{2} \left( p_i + p_j + p_k - p_ip_j + p_j + p_i + p_k - 2 \left( p_ip_j + p_j + p_k \right) \right). \tag{6}$$

While the linear term of the right-hand side (6) is independent of the pairing, the quadratic term does. $W(i,j) + W(k,b)$ is minimized when $p_i + p_j + p_k$ is maximized.

**Lemma 1.** Given $p_i \leq p_j \leq p_k \leq p_b$, the pairing $\{\text{FF}_{i,j}, \text{FF}_{k,b}\}$ is optimal.

**Proof.** It follows that $(W(i,j) + W(k,b)) - (W(i,k) + W(j,b)) = -\alpha_2 (p_i - p_j) / 2 \leq 0$. Similarly, $(W(i,j) + W(k,b)) - (W(i,k) + W(j,b)) = -\alpha_2 (p_i - p_j) / 2 \leq 0$. **■**

Let $n$ be even (odd $n$ is discussed later) and $P : \{\text{FF}_{i_1,b_1}, \ldots, \text{FF}_{i_n,b_n}\}$ be a pairing of FF$_1$, FF$_2$, FF$_3$, FF$_n$, in $n/2$ 2-bit DDCG MBFFs. The following energy waste $W(P)$ results in

$$W(P) = \sum_{i=1}^{n/2} W(i,j) = \frac{\alpha_2}{2} \sum_{i=1}^{n/2} (1 - p_i) \left( 1 - p_i \right). \tag{7}$$

It follows from (6) that $W(P)$ is minimized when $\sum_{i=1}^{n/2} p_i p_j$ is maximized. The optimal pairing could be found in polynomial time by applying a MCMC [12] to the $n$-vertex complete weighted graph, the vertices of which are FF, and its edge weights are $p_ip_j$, $1 \leq i < j \leq n$. The observation made for pairing of four FFs hints that optimal pairing should prefer FFs with close switching probabilities to reside in the same MBFF. The generalization for pairing of $n$ FFs is subsequently discussed, proving that it can be found in $O(n \log n)$ time complexity by sorting.

**Theorem 1.** Let $n$ be even and let FF$_1$, FF$_2$, FF$_3$, FF$_n$ be ordered such that $p_1 \leq p_2 \leq \cdots \leq p_n$. The pairing $P : \{\text{FF}_{i_1,b_1}, \ldots, \text{FF}_{i_n,b_n}\}$ minimizes $W(P)$ given in (7).

**Proof.** Eq. (7) shows that $W(i,j)$ is independent of the order of the FFs within a pair. It is therefore assumed w.l.o.g that for a pair $(i_1, i_2)$, there is $p_{i_1} < p_{i_2}$. By (7) $W(P)$ is independent of the order of the pairs in the summation, so $P$ is assumed w.l.o.g to be increasingly ordered such that $(i_1, i_2)$ precedes $(j_1, j_2)$ if $p_{i_1} < p_{j_1}$, $1 \leq i_1, i_2 \leq n/2$ and $i_1 \neq j_1$. Assume in contrary that there is an increasingly ordered pairing $P' : \{\text{FF}_{i_1,b_1}, \ldots, \text{FF}_{i_{n/2},b_{n/2}}\}$, $P' \neq P$, minimizing $W$. Let us compare the $n/2$ pairs of $P'$ with those of $P$ by their order, namely FF$_{(i_{2i-1},i_{2i})}$ in $P$ with FF$_{(i_{2i-1},i_{2i})}$ in $P'$, $1 \leq i \leq n/2$. Let FF$_{(2i-2,2i)}$ in $P$ and FF$_{(2i-2,2i)}$ in $P'$ be the first unmatched pairs. It follows from the increasing order of the pairs that $s_{2i-1} = 2i - 1$ and $s_{2i} > 2i$.

Consider the pair FF$_{(2i-1,2i)}$ in $P$ which follows FF$_{(2i-2,2i)}$ in $P'$. Let us derive a pairing $P''$ from $P'$ by exchanging FF$_{(2i-1,2i)}$ with FF$_{(2i,2i)}$. Assume w.l.o.g that $s_{2i} < t$. The pairing $P''$ and $P'$ thus differ on $\{\text{FF}_{(2i-1,2i)}, \text{FF}_{(2i,2i)}\} \subset P'$ and $\{\text{FF}_{(2i-1,2i)}, \text{FF}_{(2i,2i)}\} \subset P''$, and are identical on the rest $n/2 - 2$ pairs. The inequality $W(P'') - W(P') < 0$ follows from Lemma 1, thus concluding that $P$ is optimal. **■**

The time complexity of finding the optimal MBFF pairing is $O(n \log n)$ since only sorting of FFs’ toggling probabilities is required. In case of odd $n$ we could artificially add a never toggling FF, hence $p = 0$. **Theorem 1** will apply, and the optimal pairing yields FF$_{(2,3)}, \ldots, \text{FF}_{(n-1,n)}$, whereas FF$_1$ will stay unpaired.

**4. Optimal FF grouping of $k$-bit MBFF**

The hardware overhead involved in DDCG may sometimes make its application questionable for groups comprising two FFs. It has been shown in [17] that DDCG is very useful for groups of three and more FFs, depending on their toggling probabilities. We subsequently analyze the case of $k$-bit MBFFs. Let FF$_{(i_1, \ldots, i_k)}$ denote a $k$-bit MBFF comprising FF$_{i_1}, \ldots, \text{FF}_{i_k}$ and consider its energy waste $W(i_1, \ldots, i_k)$. When none of its underlying FFs is toggling, its DDCG disables the clock pulse, so energy is not wasted. Other than that DDCG enables the clock pulse. When all the FFs are toggling, the clock pulse is anyway required by all the FFs, so there is no energy waste. A waste occurs when $k - r, 1 \leq r \leq k - 1$, of the FFs are toggling, while $r$ are not. There are $\binom{k}{r}$ events of this kind and they are pairwise distinct. Since the clock pulse drives $r$ non-toggling FFs, the energy waste is $\alpha r / k$ multiplied by the probability of that event. For each $1 \leq m \leq \binom{k}{r}$ we split FF$_{i_1}, \ldots, \text{FF}_{i_k}$ into $A_m$ and $B_m$, the indices of the toggling and non-toggling FFs, respectively, $A_m \cap B_m = \{i_1, \ldots, i_k\}$, $A_m \cap B_m = \varnothing$, $|A_m| = k - r$ and $|B_m| = r$. The corresponding energy waste is therefore

$$W(i_1, \ldots, i_k) = \frac{\alpha r}{k} \sum_{r=1}^{\binom{k}{r}} \sum_{m=1}^{r} \sum_{i \in A_m} p_i \sum_{j \in B_m} (1 - p_j). \tag{8}$$

Consider $n = 2k$ FFs FF$_{i_1}, \ldots, \text{FF}_{i_{2k}}$ such that $p_{i_1} \leq p_{i_2} \leq \cdots \leq p_{i_{2k-1}} \leq p_{i_{2k}}$. We subsequently generalize **Lemma 1** for $k > 2$. It is shown that the minimum energy waste occurs for the grouping $\{i_1, \ldots, i_k\}, \{i_{k+1}, \ldots, i_{2k}\}$, namely, the $k$ FFs with the smaller probabilities in one group while the $k$ FFs with the larger probabilities in the other.

**Lemma 2.** Given $2k$ FFs FF$_{i_1}, \ldots, \text{FF}_{i_{2k}}$ satisfying $p_{i_1} \leq p_{i_2} \leq \cdots \leq p_{i_{2k-1}} \leq p_{i_{2k}}$. The grouping $\{i_1, \ldots, i_k\}, \{i_{k+1}, \ldots, i_{2k}\}$ minimizes the energy waste.

**Proof.** Assume that there exists a grouping $\{i_1, \ldots, i_k\}, \{i_{k+1}, \ldots, i_{2k}\}$ such that $W(i_{k+1}, \ldots, i_{2k}) + W(i_1, \ldots, i_k)$ is minimal, but $(i_1, \ldots, i_k) \neq (i_{k+1}, \ldots, i_{2k})$. We split the indices of the smaller and larger probabilities, $L = \{i_1, \ldots, i_k\}$ and $H = \{i_{k+1}, \ldots, i_{2k}\}$, respectively, such that $L \cup L' \cap H \cup H' \subset \varnothing$, $H' = \{i_{k+1}, \ldots, i_{2k}\} \cap \{i_1, \ldots, i_k\}$, $L'' = \{i_{k+1}, \ldots, i_{2k}\} \cap \{i_1, \ldots, i_k\}$. For every $m, 1 \leq m \leq \binom{k}{r}$, we further split $L', L'', H', H''$ according to their corresponding $k - r$ toggling FFs $A_m' \cap r$ non-toggling FFs $B_m''$ as follows:

$$L' = A_m' \cup B_m'', \quad A_m' \cap B_m'' = \varnothing;$$
$$L'' = A_m' \cup B_m'', \quad A_m' \cap B_m'' = \varnothing;$$
$$H' = A_m' \cup B_m'', \quad A_m' \cap B_m'' = \varnothing;$$
$$H'' = A_m' \cup B_m'', \quad A_m' \cap B_m'' = \varnothing. \tag{10}$$
Consider the energy wastes occurring by the grouping \(\{(i_1, \ldots, i_k), (k+1, \ldots, j_2)\}\) and the grouping \(\{(i_1, \ldots, i_k), (k+1, \ldots, j_2)\}\). Substitution of (10) in (8) yields the following expressions

\[
W_{(i_1, \ldots, i_k)} + W_{(k+1, \ldots, j_2)} = \frac{a_k}{k} \sum_{r=1}^{k-1} \sum_{m=1}^{r} \left[ \prod_{s \in A_m^0} p_s \prod_{t \in E_m^0} (1 - p_t) \prod_{s \in E_n'} (1 - p_s) \right] \cdot (d_m) \cdot (f_m) + \prod_{s \in A_m^0} p_s \prod_{t \in E_m^0} (1 - p_t) \prod_{s \in E_n'} (1 - p_s) \right]. \tag{11}
\]

\[
W_{(i_1, \ldots, i_k)} + W_{(k+1, \ldots, j_2)} = \frac{a_k}{k} \sum_{r=1}^{k-1} \sum_{m=1}^{r} \left[ \prod_{s \in A_m^0} p_s \prod_{t \in E_m^0} (1 - p_t) \prod_{s \in E_n'} (1 - p_s) \right] \cdot (d_m) \cdot (f_m) + \prod_{s \in A_m^0} p_s \prod_{t \in E_m^0} (1 - p_t) \prod_{s \in E_n'} (1 - p_s) \right]. \tag{12}
\]

The partial products in (11) and (12) are identified by appropriate symbols. Using those symbols, we have

\[
\left( W_{(i_1, \ldots, i_k)} + W_{(k+1, \ldots, j_2)} \right) - \left( W_{(1, \ldots, k)} + W_{(k+1, \ldots, j_2)} \right) = \frac{a_k}{k} \sum_{r=1}^{k-1} \sum_{m=1}^{r} \left[ (a_m e_{m} c_{n} g_{m} + b_{m} f_{m} d_{m} h_{m}) - (a_m e_{m} c_{m} g_{m} + e_{m} f_{m} g_{m} h_{m}) \right] = \frac{a_k}{k} \sum_{r=1}^{k-1} \sum_{m=1}^{r} \left( a_m e_{m} c_{m} - f_{m} h_{m} \right) \left( e_{m} g_{m} - b_{m} d_{m} \right) \leq 0. \tag{13}
\]

The inequality in (13) follows since the terms comprising the product \(a_m e_{m} c_{m}\) are all smaller than those comprising \(f_{m} h_{m}\), while the terms comprising the product \(e_{m} g_{m}\) are all larger than those comprising \(b_{m} d_{m}\). We conclude that the grouping \(\{(i_1, \ldots, i_k), (k+1, \ldots, j_2)\}\) minimizes the energy waste.

**Theorem 2.** Let \(n\) be divisible by \(k\) and let \(FF_1, FF_2, \ldots, FF_n\) be ordered such that \(p_1 > p_2 < \cdots < p_n\). The grouping \(P: \{(k+1, \ldots, j_k)\}_{i=1}^{n/k}\) minimizes \(W(P)\) given in (14).

**Proof.** Eq. (8) shows that \(W_{(s_{k+1}, \ldots, s_n)}\) is independent of the order of the FFs within a \(k\)-bit group. It is therefore assumed w.l.o.g. that there is \(s_{k+1} < s_{k+2} < \cdots < s_{2k}\). Since \(W\) in (14) is independent of the order of the terms in the summation, \(P\) is assumed w.l.o.g. to be increasingly ordered such that \((s_{k+1}, \ldots, s_n)\) precedes \((s_{k+2}, \ldots, s_n)\) if \(s_{k+1} < s_{k+2} < \cdots < s_k\), for \(1 \leq i, j \leq n/k\) and \(i \neq j\).

Assume in contrary that there is an increasingly ordered grouping \(P: \{(s_{k+1}, \ldots, s_n)\}_{i=1}^{n/k}\) \(\neq P\), minimizing \(W(P)\). Let us compare the \(n/k\) groups of \(P\) with those of \(P'\) by their order, namely \((k+1, \ldots, j_k)\) in \(P\) with \((s_{k+1}, \ldots, s_n)\) in \(P', 1 \leq i \leq n/k\). Let \((k+1, \ldots, j_k) \in P\) and \((s_{k+1}, \ldots, s_n) \in P'\) be the first unmatched groups. It follows from the increasing order of \(P\) and \(P'\) that \((s_{k+1}, \ldots, s_n) \in P'\) preceding \((s_{k+2}, \ldots, s_n) \in P\). By the increasing order of the indices in a group, \((s_{k+1}, \ldots, s_n) \in P'\) and \((s_{k+2}, \ldots, s_n) \in P\) would have been identical.

Let us derive a grouping \(P'\) from \(P\) by rearranging the \(2k\) indices of \(\{(s_{k+1}, \ldots, s_n)\}_{i=1}^{n/k}\) such that the \(k\) smaller ones reside in a group \((s_{k+1}, \ldots, s_n)\) \(\in P'\), while the \(k\) larger indices reside in a group \((s_{k+2}, \ldots, s_n) \in P'\), where the indices are renamed appropriately. Other than \(\{(s_{k+1}, \ldots, s_n)\}_{i=1}^{n/k}\) \(\subset P'\) and \((s_{k+2}, \ldots, s_n) \in P'\) are identical on the rest \(n/k-2\) groups. It then follows by Lemma 2 that \(w(P') - w(P) < 0\), which contradicts that the energy wasted by \(P'\) is minimal, thus concluding that \(P\) is optimal.

The time complexity of finding the optimal \(k\)-bit MBFF grouping is \(O(n \log n)\) since only sorting of FFs’ toggling probabilities is required. If \(n\) is not an integer multiple of \(k\), we could artificially add \(r = n - n \mod k\) never toggling FFs, and their toggling probability is therefore \(p = 0\). The problem comprising \(n+r\) FFs thus obtained obeys Theorem 2. The artificial FFs \(FF_1, FF_2, \ldots, FF_r\) will then be grouped in a \((k-r)\)-bit MBFF, whereas the rest \(n\) \(k+r\) FFs will be grouped in \(n/k\) \(k\)-bit MBFFs. It should be noted that Theorem 2 is still valid for \(p_1 < p_2 \leq \cdots \leq p_n\), where any equality can be arbitrarily resolved.

**Acknowledgments**

The authors are grateful for the corrections and useful comments and suggestions made by the reviewers.

**References**


