

Practical Boolean Decomposition for Delay-driven LUT Mapping

Alessandro Tempia Calvino¹, Alan Mishchenko², Giovanni De Micheli¹, Robert Brayton²

¹Integrated Systems Laboratory, EPFL, Lausanne, Switzerland

²Department of EECS, University of California, Berkeley, USA

Abstract—Ashenhurst-Curtis decomposition (ACD) is a decomposition technique used, in particular, to map combinational logic into *lookup tables* (LUTs) structures when synthesizing hardware designs. However, available implementations of ACD suffer from excessive complexity, search-space restrictions, and slow run time, which limit their applicability and scalability. This paper presents a novel fast and versatile technique of ACD suitable for delay optimization. We use this new formulation to compute two-level decompositions into a variable number of LUTs and enhance delay-driven LUT mapping by performing ACD on the fly. Experiments with heavily optimized benchmarks show an average delay improvement of 12.39% and an area reduction of 2.20% compared to state-of-the-art LUT mapping, with affordable run time. Additionally, our method improves the best-known delay for 4 benchmarks in the EPFL synthesis competition.

Index Terms—Logic synthesis, Boolean decomposition, technology mapping, FPGA

I. INTRODUCTION

Ashenhurst-Curtis decomposition (ACD) [1], [2], also known as Roth-Karp decomposition [3], is a powerful technique that finds a decomposition of a Boolean function into a set of sub-functions and a composition function with reduced support. ACD finds applications in logic optimization and technology mapping. The noteworthy use cases of ACD are in mapping into standard cells [4] and *field-programmable gate arrays* (FPGA) [5], decomposition of multi-valued relations [6], and encoding of multi-valued networks [7].

Traditional applications rely on the original formulation of ACD [1]–[3], breaking the input variables into two groups: the bound set (BS) and the free set (FS). Other approaches to ACD [5] allow for a shared set (SS) when one or more LUTs in terms of the BS variables are single-variable functions (buffers). The larger the SS size, the fewer LUTs are required. For instance, Figure 1 shows an ACD of a function with BS, FS, and SS resulting in three 5-input LUTs. In [5], maximizing the SS is implemented using *binary decision diagrams* (BDDs) [8]. More recently, truth-table-based implementations eliminated the need for explicitly constructing a BDD, resulting in a faster decomposition [9], [10].

ACD has been applied to map into fixed *lookup table* (LUT) structures [10] as a way to mitigate structural bias and improve the quality of standard LUT mapping. This approach utilizes heuristic variable re-ordering to find an ACD, supporting up to 1 SS variable. Additionally, ACD has been used in post-mapping resynthesis [9], when logic cones composed of several LUTs are collapsed into single-output Boolean

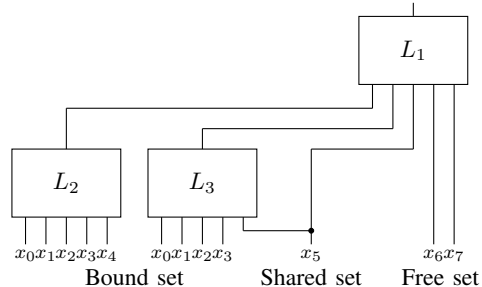


Fig. 1. ACD of an 8-input Boolean function into three 5-input LUTs with a 5-variable *bound set* (BS), a 1-variable *shared set* (SS), and a 2-variable *free set* (FS).

functions and re-expressed using fewer LUTs. The authors proposed to use *disjoint-support* decomposition (DSD) and Shannon’s expansion to pack logic into LUTs while supporting up to 3 SS variables.

Since ACD is often applied only to functions up to 11 or 16 inputs (for LUT structures composed of two or three 6-LUTs, respectively), state-of-the-art LUT mapping is performed through local substitutions applied to an initial graph representation, called *subject graph*. Generally, delay-optimal mapping w.r.t. the subject graph is feasible in polynomial time [11], while area-optimal mapping is NP-hard [12]. However, the structure of the subject graph highly impacts the result. This phenomenon is known as *structural bias*. To mitigate structural bias, methods in the literature generate a set of structural choices (or decompositions) available during mapping [13]–[15].

This paper offers two main contributions. First, we revisit the formulation of ACD with SS to enhance its computationally efficiency in LUT mappers and post-mapping resynthesis engines performing delay optimization. Our algorithm is truth-table-based and flexible in the number of FS, BS, and SS variables, and in the number of BS functions. Our ACD runs up to 2x faster, compared to [10], and up to 80x faster, compared to [9] when performing decompositions into two 6-LUTs. Furthermore, it also finds considerably more solutions.

Second, we use ACD for the delay optimization of LUT networks. The idea is to compute functional decompositions using the timing-critical variables in the FS and the rest of the variables in the BS and SS. We integrate our ACD into the state-of-the-art LUT mapper for delay optimization. To our knowledge, this is the first practical and scalable work that

uses ACD for delay-driven LUT mapping.

We experimentally evaluate the performance of ACD and compare mapping based on Boolean decomposition against state-of-the-art methods:

- 1) We compare our ACD method against other decomposition methods in ABC, showing better quality with a competitive or better run time.
- 2) We demonstrate that mapping with ACD can efficiently mitigate structural bias and considerably reduce the delay. We compare the default LUT mapper in ABC, the LUT mapper with Boolean decomposition in ABC, and the proposed mapper with integrated ACD. We show that mapping with ACD outperforms the other mappers in delay by 7.52% on average with and without structural choices [15]. Moreover, we show that an additional mapping round using the network obtained by ACD as a structural choice can further improve the delay, compared to the standard LUT mapper, by 12.39% with an area reduction of 2.20%.
- 3) We present 4 new best results in the EPFL competition.

II. PRELIMINARIES

This section introduces the basic notations and background related to logic networks, decomposition, and LUT mapping.

A. Definitions

A *Boolean function* is a mapping from a k -dimensional Boolean space into a 1-dimensional one: $\{0, 1\}^k \rightarrow \{0, 1\}$.

A *truth table* representation of a k -input Boolean function $f : \{0, 1\}^k \rightarrow \{0, 1\}$ can be encoded as a bit string $b = b_{l-1} \dots b_0$, i.e., a sequence of bits, of length $l = 2^k$. A bit $b_i \in \{0, 1\}$ at position $0 \leq i < l$ is equal to the value taken by f under the input assignment $\vec{a} = (a_0, \dots, a_{k-1})$ where

$$2^{k-1} \cdot a_{k-1} + \dots + 2^0 \cdot a_0 = i.$$

The *positive cofactor* of a Boolean function f with respect to a variable x_i , represented as f_{x_i} , is the Boolean function obtained by setting $x_i = 1$. Similarly, the *negative cofactor* \bar{f}_{x_i} is the Boolean function obtained by setting $x_i = 0$.

In the classical representation, we refer to the leftmost input column of a truth table as the *most significant variable* (a_{k-1}) and the rightmost input column as the *least significant variable* (a_0). A *swap* of two variables results in the interchange of the corresponding two-variable cofactors, thereby altering the truth table.

Figure 2 depicts two truth tables represented as bit strings, one in binary and one in hexadecimal. Notably, the rightmost truth table can be derived from the leftmost one by swapping the variables x_0 and x_2 . Marked next to both truth tables are the cofactors with respect to two most significant variables.

A completely specified Boolean function f *essentially depends* on a variable v if there exists an input combination such that the value of the function changes when the variable is toggled ($\frac{\partial f}{\partial v} = 1$). The *support* of f is the set of all variables on which function f essentially depends. The supports of two functions are *disjoint* if they do not contain common variables.

x_2	x_1	x_0	f	$x_0 \leftrightarrow x_2$	x_0	x_1	x_2	f
0	0	0	1		0	0	0	1
0	0	1	0		0	0	1	1
0	1	0	1		0	1	0	1
0	1	1	0		0	1	1	0
1	0	0	1		1	0	0	0
1	0	1	1		1	0	1	1
1	1	0	0		1	1	0	0
1	1	1	1		1	1	1	1
$f = 10110101$					$f = 0xA7$			

Fig. 2. Truth table representations and their encoding, cofactor extraction w.r.t. the two most significant variables, and variable swapping of x_0 with x_2 .

A set of functions is disjoint if their supports are pair-wise disjoint.

A *Boolean network* is modeled as a directed acyclic graph (DAG) with nodes represented by Boolean functions. The sources of the graph are the *primary inputs* (PIs), the sinks are the *primary outputs* (POs). For any node n , the *fanins* of n is a set of nodes driving n , i.e. nodes that have an outgoing edge towards n . Similarly, the *fanouts* of n is a set of nodes driven by node n , i.e., nodes that have an incoming edge from n . A *k-LUT network* is a Boolean network composed of k -input *lookup tables* (k -LUTs) capable of realizing any k -input Boolean function. An *and-inverter graph* (AIG) [16] is a Boolean network where nodes are 2-input ANDs and edges may implement inverters.

A *cut* C of a Boolean network is a pair (n, \mathcal{K}) , where n is a node called *root*, and \mathcal{K} is a set of nodes, called *leaves*, such that 1) every path from any PI to node n passes through at least one leaf and 2) for each leaf $v \in \mathcal{K}$, there is at least one path from a PI to n passing through v and not through another leaf. The *size* of a cut is the number of leaves. A cut is k -feasible if its size does not exceed k .

B. Ashenhurst-Curtis decomposition

Ashenhurst-Curtis decomposition (ACD) [1]–[3], of a single-output Boolean function f can be expressed as follows:

$$f(\vec{x}_{bs}, \vec{x}_{ss}, \vec{x}_{fs}) = g(\vec{h}(\vec{x}_{bs}, \vec{x}_{ss}), \vec{x}_{ss}, \vec{x}_{fs}), \quad (1)$$

where \vec{x}_{bs} is the *bound set* (BS), \vec{x}_{ss} is *shared set* (SS), and \vec{x}_{fs} is the *free set* (FS). These sets are disjoint variable subsets, which together form the support of f . The function \vec{h} may be multi-output with the number of outputs less than the BS size. The single-output functions in \vec{h} are referred to as BS functions. The function g is referred to as the *composition function*. When decomposing into k -LUTs, the composition function is typically chosen to fit into one k -input LUT. Figure 1 shows an ACD of an 8-input function into three 5-input LUTs with a 5-variable BS, a 1-variable SS, and a 2-variable FS. The decomposition generates two BS functions (L_2, L_3) and a composition function (L_1) with 5 inputs.

C. FPGA technology mapping

LUT mapping is the process of expressing a Boolean network in terms of k -input lookup tables (k -LUTs). Before

mapping, the network is represented as a k -bounded Boolean network called the *subject graph*, which contains nodes with a maximum fanin size of k . The AIG is the most common subject graph representation. The subject graph is transformed into a mapped network by applying local substitutions to sections of the circuit defined by cuts, which are computed using cut enumeration [17]. A LUT mapper computes a mapping solution by selecting a subset of the cuts that cover the subject graph while minimizing a cost function. The state-of-the-art LUT mapper computes cuts and refines the mapping solution in several mapping passes using heuristics based on delay, area, and edge count. For further details, refer to [18].

III. IMPROVEMENTS TO ACD

This section discusses a fast and versatile truth-table-based implementation of ACD for single-output functions with support for a shared set. We propose several novelties that make ACD practical within LUT mappers and resynthesis methods. Figure 3 illustrates the ACD computation. The BS, SS, FS, and the number of BS functions used are flexible and determined during the decomposition. The composition function (L_1) is implemented as a multiplexer of cofactors with respect to BS functions and the shared set. Functions dependent on the FS (g_i), called FS functions, are the data inputs of the multiplexer found inside the composition function. BS functions and the shared set are instead the selection inputs.

This definition of decomposition reflects the one used by previous approaches [5]. Specifically, the decomposition is generic, i.e., it includes other types of decomposition. For instance, a Shannon's expansion:

$$f = x f_x + \bar{x} f_{\bar{x}},$$

where x is a selector of a multiplexer, can be re-expressed in our ACD format:

$$f = f_x f_{\bar{x}} 1 + f_x \bar{f}_{\bar{x}} x + \bar{f}_x f_{\bar{x}} \bar{x} + \bar{f}_x \bar{f}_{\bar{x}} 0,$$

where x is a FS variable, f_x and $f_{\bar{x}}$ are BS functions, and FS functions g_i are 1, x , \bar{x} , and 0.

In this section, we first present how to efficiently check the existence of a feasible ACD and assign variables to the FS, BS, and SS (Section III-A). Next, we show how to compute the decomposition while minimizing the number of BS functions and their support (Section III-B).

A. Finding a feasible decomposition

After defining the properties of ACD, in this section we present an efficient method to check the existence of a Boolean decomposition and find an assignment of support variables to the FS and the BS (and SS). In particular, we focus on decomposition into two levels of k -input LUTs. For simplicity, in this section we consider SS variables a part of the BS.

The first step to derive a decomposition is to partition of variables into FS and BS. Given a truth table, our approach enumerates different free sets. Let N be the number of variables in the support of a function to decompose. Let P be the number of variables to consider in the FS. The remaining

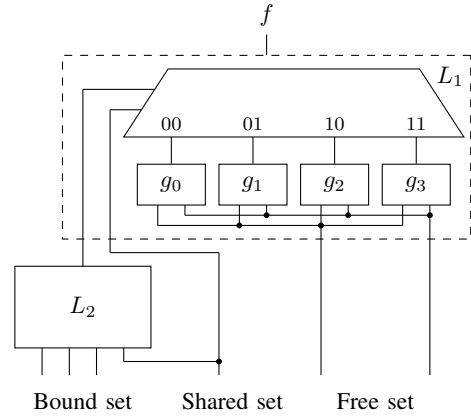


Fig. 3. Illustrating the AC decomposition of a Boolean function

$N - P$ variables are considered in the BS. The number of different free sets is $\binom{N}{P}$. Regarding the choice of value P when searching for a feasible two-level decomposition, for an N -input function and k -input LUTs, it is convenient to consider $(N - k)$ variables in the FS, so that at most k variables are considered in the BS. For instance, when $N = 8$ and $k = 6$, we can choose $P = 2$ and evaluate $8 \cdot 7/2 = 28$ different 2-variable free sets.

For each FS, the truth table is transformed to have the FS variables as the least significant ones, compared to the BS variables. The variable reordering is performed using a dedicated procedure, which swaps two variables. Note that when enumerating all the free sets the first FS composed of the P least significant variables in the support of the function does not need variable swapping since the original truth table already reflects this order. Then, every consecutive FS can be derived from a previous FS by swapping one variable in x_{fs} with one in x_{bs} . The complexity to explore all the FS is of $\binom{N}{P}$ swap operations. Figure 2 shows how a variable swap affects the truth table.

Each input assignment to the BS variables selects one P -input function in terms of the FS variables. Specifically, each P -input function is a cofactor with respect to x_{bs} . From a truth table in this format, FS functions are easily computed by extracting groups of 2^P bits at $i \cdot 2^P$ offsets with $i \in [0, 2^{(N-P)})$. Informally, FS functions are listed next to each other. Figure 2 graphically depicts the extraction of cofactors with respect to the two most significant variables.

Example 1: Let us consider the 6-variable function represented in hexadecimal format as a truth table $f = 0x8804800184148111$. Let us assume that the FS variables are the two least significant variables and the BS variables are the four most significant ones. The functions in terms of FS variables have truth tables with $2^P = 2^2 = 4$ bits. There are $2^{(N-P)} = 16$ of them, corresponding to hexadecimal digits in the truth table (0x8, 0x8, 0x0, 0x4, etc). \triangle

The target function can be realized using M bound set functions if the number of unique FS functions, known as column multiplicity μ , does not exceed 2^M , hence $M \geq \lceil \log_2(\mu) \rceil$. If

$P + M \leq k$, the composition function can be implemented as a k -LUT.

Example 2: Continuing Example 1, there are 16 FS functions of which only 4 are unique. The FS functions are 0x8, 0x0, 0x4, and 0x1. Hence, the column multiplicity $\mu = 4$, which needs at least $M = \lceil \log_2(4) \rceil = 2$ BS functions. Hence, this partition of variables into FS and BS produces a feasible support-reducing decomposition into 4-input LUTs. Using Figure 3, ACD assigns FS functions to g_i . Then, two BS functions of at most 4 inputs are necessary to select the correct FS function. \triangle

We employ the enumeration of free sets while counting the number of unique cofactors to check if a support-reducing decomposition exists. In practice, a sufficient condition for a 2-level decomposition is to have $M + P \leq k$ and $N - P \leq k$, i.e., the composition function is k -feasible, and the number of remaining variables in the BS does not exceed k .

After identifying a partition of variables into FS and BS, and the corresponding unique FS functions, our method uses the techniques in Section III-B to produce a decomposition while minimizing the number of BS functions and their support.

B. Functional encoding and support minimization

Once a partition of variables into FS and BS with a feasible decomposition is found, the BS functions are extracted by assigning each FS function to an encoding. Informally, an encoding represents the assignment of FS functions to the data inputs of the MUX of Figure 3 (e.g., the encoding of g_1 is 01). While any encoding that distinguishes FS functions is a valid solution, a good encoding also minimizes the number of BS functions required (by maximizing the shared set), and the functional support. In particular, it is crucial to find an encoding that minimizes the support for three reasons. First, if $N - P > k$, by minimizing the support, each BS function would ideally fit into a k -LUT, and the decomposition is feasible in 2 levels. Second, minimizing the support maximizes the shared set (buffer BS functions), reducing the number of required LUTs. Third, the number of edges required is reduced, helping routability. Finding a feasible encoding is similar to solving constrained encoding problems [19]–[21].

An encoding is an assignment of a code $T = t_{M-1} \dots t_0$ of length M to each FS function. A variable t_i takes one of the three values, 1, 0, or –, indicating the ON-set, OFF-set, and DC-set, respectively. Let i -sets be the set of μ Boolean functions in terms of the BS variables encoding FS functions using one-hot encoding. Precisely, an i -set represents one FS function and takes value 1 when an input assignment to the BS variables results in the corresponding FS function.

Example 3: Using Example 2, the i -set corresponding to the FS function 0x8 is 1100100010001000 in binary format. Note that the truth table has $N - P$ variables and has value 1 when the original function is 0x8. \triangle

i -sets are used to derive a more compact encoding with a two-step procedure. The first one enumerates *candidate BS functions*. The second one solves a unate covering problem in which columns are candidate BS functions and rows are pairs of FS functions to be distinguished.

Candidate BS functions are functions depending on BS variables whose output can be used as t_i to encode FS functions. They are enumerated by combining i -sets. To leverage all the functional degrees of freedom of a strict encoding, i -sets in a BS candidate can be either in the *ON-set*, *OFF-set*, or *don't-care* (DC) set. Since candidate BSs are used as select inputs of a multiplexer, BS candidates can distinguish elements in the ON-set (takes value 1) against elements in the OFF-set (takes value 0). In encoding problems, BS functions are called *dichotomies*, while the pairs of functions to be distinguished are referred to as *seed dichotomies* [21]. Don't-cares in BS candidates are also important to minimize the support, which translates into fewer LUT edges.

Example 4: Continuing Example 3, let us consider the candidate bound set function h that has the i -sets {0x8, 0x1} in the ON-set and the i -set {0x4} in the OFF-set. Its function in binary format is $h = 11-01-110101111$ where “–” is a don't care. When $h = 1$, either 0x8 or 0x1 are selected. When $h = 0$, 0x4 is selected. The corresponding dichotomy is $\{\{0x8, 0x1\}, \{0x4\}\}$. In this case, function h distinguishes 0x8 from 0x4 and 0x1 from 0x4, covering the two seed dichotomies $\{\{0x8\}, \{0x4\}\}$ (or $\{\{0x4\}, \{0x8\}\}$) and $\{\{0x1\}, \{0x4\}\}$ (or $\{\{0x4\}, \{0x1\}\}$). \triangle

A candidate bound set function is generated by assigning each i -set to be in the ON-set, OFF-set, or DC-set. Hence, the total number of possible BS candidates is 3^μ . Nonetheless, some BS candidates are interchangeable, i.e., one candidate can be obtained by swapping the ON-set and the OFF-set of another BS candidate. Our enumeration removes these symmetries by fixing one i -set to be only in the ON-set or DC-set, enumerating only $2 \cdot 3^{\mu-1}$ BS candidates. Moreover, candidates not distinguishing any pair of FS functions are removed. As a special case, if μ is a power of 2, the number of possible BS candidates reduces to $\binom{M}{M/2}/2$ by splitting the FS functions to be equally distributed between ON-set and OFF-set, i.e., each BS candidate must distinguish half of the FS functions against the other half.

One limitation of this method is that the number of BS candidates is exponentially dependent on the column multiplicity. However, we may further reduce the number of BS candidates when it is too large. In particular, for an ACD into 6-LUTs the maximum column multiplicity to support is 16. Consequently, the highest number of BS candidates is 9.5 million for $\mu = 15$. To maintain a reasonable number of BS candidates, our method does not use don't cares for problems with $\mu > 8$, enumerating $2^{\mu-1}$ candidates and reducing the highest number of candidates to 16 thousand. Through experimentation, we have observed that imposing this limitation scarcely affects the quality of the encoding, while substantially enhancing run-time efficiency. Conversely, extending this method to lower multiplicity values noticeably compromises the solution quality.

Each BS candidate function is associated with a cost that depends on the number of variables in its support. The number of variables is computed with a special procedure that considers don't cares. Then, a covering table is constructed by

	4	3	3
	C9AF	1177	2727
$\{\{0x8\}, \{0x0\}\}$	1	0	1
$\{\{0x8\}, \{0x4\}\}$	1	1	0
$\{\{0x8\}, \{0x1\}\}$	0	1	1
$\{\{0x0\}, \{0x4\}\}$	0	1	1
$\{\{0x0\}, \{0x1\}\}$	1	1	0
$\{\{0x4\}, \{0x1\}\}$	1	0	1

Fig. 4. Covering table to solve the encoding problem.

having all the pairs of FS functions to be distinguished (seed dichotomies) as rows and the BS candidates as columns. A row-column entry (i, j) is 1 if the BS candidate of column j distinguishes the seed dichotomy i . A solution that minimizes the support is computed by solving a minimum-cost covering problem [21]. The solution must cover all the rows while minimizing the cost. We use greedy covering followed by local search to compute cost-minimizing cover. A single iteration of greedy covering extracts one column covering the most non-covered rows while minimizing the cost. The process is iterated until a solution is found. Then, the solution is iteratively improved by replacing one column with another having a lower cost.

Example 5: Figure 4 shows a covering table reflecting the examples in this section. Each column in the table is a candidate BS function shown as a truth table in hexadecimal format on 4 variables. Each BS candidate has a cost based on the number of variables on its support. Each row is a seed dichotomy. An element (i, j) in the table is 1 if the BS_j distinguishes the seed dichotomy i . The best solution with cost 6 takes the second and third columns and results in two BS functions depending on 3 variables. \triangle

Given a solution, an encoding of the FS functions is obtained by assigning a code $T = t_{M-1} \dots t_0$, in which each variable t_i corresponds to a selected BS_i candidate.

Example 6: Continuing Example 5, a minimum cover involves $BS_0 = 0x1177$, by taking $0x4$ and $0x1$ in the ON-set, and $BS_1 = 0x2727$ by taking $0x0$ and $0x1$ in the ON-set. Given the BS functions, the encoding of the FS functions assigns the following codes to g_i in Figure 3: $T_{0x8} = 00$, $T_{0x4} = 01$, $T_{0x0} = 10$, and $T_{0x1} = 11$. Finally, the composition function is computed using the FS and its encoding, resulting in function $0x1048$ when represented in hexadecimal format. Consequently, the function has been successfully decomposed using three 4-LUTs. \triangle

IV. TECHNOLOGY MAPPING WITH ACD

In this section, we leverage the Ashenhurst-Curtis decomposition (ACD) methods described in Section III to improve the delay of LUT networks. ACD can be used in two ways: 1) as part of LUT mapping or 2) as a post-mapping resynthesis method to compact logic and decrease the delay. In this work, we focus on the former usage since it has more flexibility and optimization opportunities. Although post-mapping resynthesis is not covered in this work, its implementation

would follow a methodology similar to [9]. First, this section discusses how to perform delay-oriented functional decomposition for any number of FS variables and BS functions. Then, it describes the integration of ACD in a technology mapper.

A. Delay-oriented ACD

Let us consider a node n in a k -LUT network and a cut C rooted in n that contains leaves in the input sub-network of n . Among all the leaves, some are timing-critical and some are not. Let D be the latest arrival delay of a leaf in C . We use ACD to find an implementation that realizes the function of cut C with delay $D + 1$ where $|C| > k$, assuming a unit-delay model. Specifically, we use the timing-critical leaves of C in the FS and other non-critical ones in the BS or SS. This transformation may reduce the worst delay of a LUT network when applied on the critical path.

The ACD-based transformation is performed in two steps. First, our method verifies the existence of a delay-minimizing decomposition. Second, if a decomposition exists, it solves the encoding problem and returns a solution.

1) *Checking the existence of a decomposition:* Algorithm 1 shows the procedure *evaluate* to check the existence of an ACD. The algorithm receives the function represented as a truth table tt of a large cut with size N where $N > k$. Set S contains a list of timing-critical variables with delay D . First, the truth table is transformed to have critical variables as the least significant ones since they must be in the FS (at line 3). The proposed approach limits $N - P \leq k$ to ensure a two-level decomposition without solving the encoding problem. Hence, the number of variables in the FS must be at least $P \geq N - k$, and $P \geq |S|$ to include all the delay-critical variables (at line 6). For each FS of P_i variables, the column multiplicity value is computed using the method described in Section III-A, and the smallest one is returned (at line 7). In this case, since delay-critical variables are always part of the FS, $\binom{N}{P_i - |S|}$ different combinations are enumerated. If the smallest multiplicity found can be implemented using at most $k - P_i$ BS functions, a delay-minimizing ACD exists. In this case, variables in the FS have the delay increase of 1 while other variables have the delay increase of 2 (at line 14). If, on the other hand, a decomposition with P_i does not exist, the function is not decomposable.

The loop in line 6 begins checking the existence of a decomposition with a smaller value of P . This approach is based on the theoretical property that if a function is not decomposable for the given value of P , it is also not decomposable for $P + 1$. Then, if a decomposition exists, the loop attempts to increase the number of variables in the free set. Specifically, maximizing the free set to include non-critical variables has multiple benefits. Primarily, the decomposition would have a reduced column multiplicity, which simplifies the encoding problem. Additionally, maximizing the free set may increase the required time of the associated non-critical signals, facilitating the area-recovery process of technology mapping.

Algorithm 1: ACD evaluation

```
1 Input : Truth table  $tt$ , LUT size  $k$ , Late vars set  $S$ 
2 Output: Propagation delay
3  $\text{reorder\_variables}(tt, S)$ ;
4  $\mu_{best} \leftarrow \infty$ ;
5  $x_{fs} \leftarrow \emptyset$ ;
6 for  $P_i \leftarrow \max(\text{num\_vars}(tt) - k, |S|)$  to  $k - 1$  do
7    $\{\mu, x'_{fs}\} \leftarrow \text{compute\_smallest\_multiplicity}(tt, P_i, |S|)$ ;
8   if  $\mu \leq 2^{k-P_i}$  and  $\mu < \mu_{best}$  then
9      $\mu_{best} \leftarrow \mu$ ;
10     $x_{fs} \leftarrow x'_{fs}$ ;
11    continue;
12 break;
13 if  $\mu_{best} \neq \infty$  then
14   return  $\text{compute\_propagation\_delay}(tt, x_{fs})$ ;
15 return  $\text{infinite\_propagation\_delay}()$ ;
```

2) *Computing the decomposition*: After applying *evaluate*, another procedure *decompose* is used to compute the actual decomposition using the methods described in Section III-B.

B. LUT mapping with ACD

The methods described in Section IV-A have been integrated into the LUT mapping algorithm in [18]. Each mapping iteration computes k -feasible cuts rooted in nodes of the subject graphs and selects one best cut for each node based on cost functions and slack. Typically, enumerated cuts are k -feasible, i.e., any cut abstracts a k -LUT. In our implementation, cut enumeration computes large cuts up to size $k < l \leq 11$, where l is provided by the user. During cut enumeration, the mapper computes cut functions as truth tables. For the non- k -feasible computed cuts, the mapper uses Algorithm 1 to check the existence of a delay-minimizing decomposition into k -LUTs. If a decomposition is not feasible, the cut is discarded. If a decomposition exists, the cut delay is computed using the propagation delay returned by Algorithm 1. The area is computed pessimistically, neglecting the existence of a shared set, i.e., $\text{Area} = \lceil \log_2 \mu \rceil + 1$. To have precise area information, i.e., the number of required LUTs, ACD would need to solve the encoding problem and compute the decomposition. However, experimentally, not running the decomposition on the fly reduces the run time considerably with negligible impact on the final circuit area.

The mapper uses l -feasible cuts with ACD in the delay mapping pass, while it uses k -feasible cuts in the following area recovery iterations. Note that area-recovery aims at improving the solution over non-critical paths and can always re-use the best cuts from the previous pass, such that the required times are met. After the last mapping pass, a cover is generated consisting of k - and l -feasible cuts. At this stage, the mapper decomposes the non- k -feasible cuts into k -LUTs.

V. EXPERIMENTS

This section presents an experimental evaluation of the proposed LUT mapping with ACD. First, the ACD algorithm

proposed in this paper is compared with other state-of-the-art methods for decomposing practical functions. Then, we evaluate ACD for delay-driven LUT mapping. While the experiments are reported for 6-input LUTs, similar improvements have been obtained for 4-input LUTs as well.

The proposed methods have been implemented in *ABC* [22]. For our experiments, we use the EPFL combinational benchmark suite [23] containing several circuits provided as *and-inverter graphs* (AIGs). The baseline has been obtained using the commands and scripts “*dfraig; resyn; resyn2; resyn2rs; if -y -K 6; resyn2rs*” in *ABC*, which perform a high-effort size and depth AIG optimization. In particular, it combines SAT sweeping [24], scripts for delay-oriented AIG optimization [16], and lazy man’s logic synthesis [25], which is the most aggressive depth minimization command in *ABC*. The experiments have been conducted on an Intel i5 quad-core 2GHz on MacOS. The results have been verified using combinational equivalent checking in *ABC*. We extended the LUT mapper *if* in *ABC* to perform ACD as discussed in Section IV. The following commands are used in the experiments:

- *dch (-f)*: computes structural choices used to mitigate the structural bias [15], where *-f* stands for “fast”;
- *if -K 6*: performs delay-oriented technology mapping with choices into 6-LUTs using 6-feasible cuts;
- *if -s -S 66 -K 8*: performs delay-oriented technology mapping using 8-feasible cuts and decomposes logic for minimal delay into two 6-LUTs using a SAT-based formulation (available in *ABC* but not published);
- *if -Z 6 -K 8*: performs technology mapping into 6-LUTs using the proposed implementation of delay-oriented ACD described in Section IV for 8-feasible cuts;
- *st*: derives an AIG from an LUT network.

A. Decomposition success rate

In this experiment, we evaluate the performance of ACD in decomposing functions by comparing it against other implementations in *ABC*. Specifically, we test the number of functions that can be successfully decomposed into two 6-LUTs and the run time needed. We run this experiment on *practical functions*, i.e., functions that are observable in designs and benchmarks, which include fully-, partially-, and non-DSD-decomposable functions. We extract practical functions from the EPFL benchmarks. Since the number of practical functions can be large, we classify them into \mathcal{NPN} -equivalence classes employing the heuristic sifting algorithm [26], [27].

Table I shows the percentage of decomposable functions and the runtime for different methods and support sizes. For instance, the first column contains results for decomposing practical 7-input functions, where (41071) indicates the number of unique NPN classes collected. Each row of the table shows one ACD method. The first method *lutpack* [9] performs a heuristic ACD using DSD and the Shannon’s expansion, supporting up to 3-SS variables. The second method, *S66* [10], performs ACD using heuristic variable re-ordering supporting at most 1-SS variable. Finally, we present two variants of our

TABLE I
DECOMPOSITION SUCCESS RATIO INTO TWO 6-LUTs FOR PRACTICAL FUNCTIONS USING DIFFERENT ACD METHODS.

ACD type	7 vars (41071)		8 vars (107466)		9 vars (195602)		10 vars (313649)		11 vars (404991)	
	Success (%)	Time(s)	Success (%)	Time(s)	Success (%)	Time(s)	Success (%)	Time(s)	Success (%)	Time(s)
lutpack [9]	98.34%	20.39	83.47%	64.37	69.92%	154.38	48.95%	334.79	26.87%	897.55
S66 [10]	84.18%	0.60	69.24%	2.57	52.13%	4.99	37.36%	6.99	19.14%	9.79
66 1-SS	97.30%	0.28	82.23%	1.41	74.24%	4.20	63.06%	9.39	32.88%	16.43
66 M-SS	99.82%	0.30	92.94%	3.08	84.71%	9.92	63.06%	9.73	32.88%	16.58

TABLE II
SUCCESS RATIO WHEN DECOMPOSING PRACTICAL FUNCTIONS INTO 2 LEVELS OF 6-LUTs WITH THE GIVEN LATE-ARRIVING VARIABLES.

N late	ACD type	7 vars	8 vars	9 vars	10 vars	11 vars
0	66 M-SS	99.82%	92.94%	84.71%	63.06%	32.88%
	Generic	100.00%	100.00%	98.05%	90.20%	32.88%
1	66 M-SS	96.59%	79.60%	61.51%	37.35%	16.54%
	Generic	100.00%	100.00%	97.57%	83.23%	16.54%
2	66 M-SS	86.22%	59.78%	39.28%	23.74%	10.95%
	Generic	100.00%	100.00%	94.19%	66.56%	10.95%
3	66 M-SS	65.11%	36.37%	21.25%	13.78%	6.96%
	Generic	93.78%	86.03%	76.82%	44.51%	6.96%
4	66 M-SS	36.96%	17.00%	8.62%	7.21%	4.43%
	Generic	54.55%	40.42%	25.45%	23.70%	4.43%
5	66 M-SS	14.52%	5.42%	2.96%	2.84%	2.61%
	Generic	14.52%	5.42%	2.96%	2.84%	2.61%

decomposition method restricted to use 2 6-LUTs. One uses up to 1-SS variable (66 1-SS), the other (66 M-SS) has no restrictions on the number of SS variables. The approaches described in this paper outperform the state of the art in quality for a competitive or better run time.

B. Decomposition success rate for delay optimization

We extend the previous experiment to evaluate delay minimization using the proposed ACD method. This experiment tests the success rate of the decomposition for practical functions given delay-critical variables, which are required to be in the free set. Informally, for delay-critical variables with delay D , this experiment checks the existence of a decomposition with delay $D+1$. We only consider 66 M-SS and generic ACD since other known methods do not perform delay minimization using the input arrival time. For each function, we randomly generate up to 10 unique sets of delay-critical variables and test the decomposition for each one of them. While 66 M-SS is limited to two LUTs, generic can use up to 4 LUTs.

Table II presents the success rate based on the number of delay-critical variables, shown in the column “N late”. The table highlights the advantage of supporting multiple BS functions. Generic ACD has a high success rate in most cases. Limitations occur when the number of delay-critical variables exceeds 3 or the number of variables in the support is 10 or more. Generally, the decomposition of 11-input variables is rare. However, many 10 input variables are still decomposable.

C. Delay-driven LUT mapping

Table III compares four technology mapping strategies for delay minimization during mapping into 6-LUTs, assuming a unit-delay model. Each strategy takes the baseline as an input and computes structural choices before mapping. Structural choices have not been used for the benchmark *hyp* due to a known bug in ABC. The proposed method is compared against standard LUT mapping and mapping into LUT structures. Command *ACD* denotes our mapper with Boolean decomposition using the sequence “dch; if -Z 6 -K 8”. We do not compare against [10] and [9] because those methods do not support delay minimization. Furthermore, we do not compare against the recent mapper with gate decomposition based on bin-backing [28]. Nevertheless, the mapper in [28] would improve the average delay of ABC *if* by only 0.31%.

Mapping into LUT structure “66” composed of two 6-LUTs, which is a SAT-based version of structural ACD, reduces depth by 1.04% and the area by 2.57% on average, at the cost of increasing the number of edges by 2.57%. The proposed LUT mapping with ACD improves the depth of the LUT network by 7.52% on average while increasing the number of LUTs and edges by 8.13% and 7.87%, respectively.

Note that most of the improvement is concentrated in the first 10 benchmarks since others are already close to their best known depth [29]. For 4 of them, the delay reduction exceeds 20% and is up to 27.27%. Practically, part of the area increase can be reduced by area-recovery methods [9], [30], [31], using delay relaxation, or by an additional mapping step applied after ACD. The rightmost strategy performs the latter option. The LUT count and edge count are reduced considerably, leading to an area improvement of 2.20%, compared to traditional technology mapping with choices. Also, the logical depth further decreases up to 54.55%. Specifically, the result after ACD is used as a choice to improve the next round of technology mapping because choices extracted from mapping with ACD are more structurally suited to delay-oriented mapping, compared to the original AIG. Moreover, structural choices help reduce the area over the non-critical paths. Note that a second mapping round does not provide practical benefits if applied to the default LUT mapper (leftmost column) since the network after deriving the AIG is structurally similar to the baseline. Furthermore, benchmark *hyp* is noticeably improved by remapping both in area and delay without using structural choices. Regarding the run time, mapping with ACD is faster than mapping into LUT structures while being more general.

TABLE III
COMPARISON OF DELAY-DRIVEN LUT MAPPING, LUT MAPPING INTO LUT STRUCTURE “66”, AND LUT MAPPING USING ACD.

Benchmark	ABC: dch; if -K 6				ABC: dch; if -s -S 66 -K 8				ACD				ACD; st; dch -f; if -K 6			
	LUTs	Edges	Depth	Time (s)	LUTs	Edges	Depth	Time (s)	LUTs	Edges	Depth	Time (s)	LUTs	Edges	Depth	Time (s)
adder	363	1433	22	0.18	362	1465	20	0.28	383	1519	16	0.20	353	1518	10	0.39
bar	1664	9344	4	0.44	1664	9344	4	0.57	1664	9344	4	0.47	1006	5274	4	0.76
div	8618	32394	406	6.62	9107	33665	397	13.42	11644	44496	326	7.16	9068	39167	271	21.19
hyp	58393	239097	1864	5.43	61701	247699	1840	31.82	65615	264998	1396	11.13	61769	263254	1034	19.76
log2	9712	43562	58	17.05	10172	44943	58	30.06	10313	46365	56	17.81	9429	42533	57	39.09
max	831	3804	14	0.37	840	3668	14	0.63	1211	5578	12	0.42	871	4277	11	1.39
multiplier	7383	34137	36	6.01	7334	32781	36	12.11	7693	35798	33	6.82	6800	31705	31	13.32
sin	1928	8445	30	1.31	1948	8463	30	4.94	2052	8913	29	1.50	1830	8178	30	2.91
sqrt	7515	29573	663	4.17	7972	30610	638	12.66	10156	38558	519	4.73	9292	36030	476	8.77
square	4122	17319	23	1.98	4165	17547	22	3.91	4107	17924	18	2.22	4118	18285	14	5.15
arbiter	1833	8982	6	1.64	1879	8836	6	2.02	1850	8987	6	1.70	2037	8780	6	3.33
cavlc	137	707	4	0.13	104	491	4	0.56	137	707	4	0.15	123	655	4	0.20
ctrl	30	133	2	0.07	28	127	2	0.08	30	133	2	0.08	29	126	2	0.08
dec	287	684	2	0.09	287	1404	2	0.1	287	684	2	0.10	284	816	2	0.12
i2c	312	1360	3	0.16	306	1316	3	0.36	319	1378	3	0.19	297	1329	3	0.27
int2float	52	258	3	0.08	46	205	3	0.18	52	258	3	0.09	50	251	3	0.11
mem_ctrl	11037	48812	18	10.24	10830	46368	18	31.67	11232	49483	17	11.40	10398	45793	16	20.57
priority	178	725	6	0.11	182	736	6	0.18	185	736	6	0.12	171	698	6	0.17
router	89	285	4	0.09	61	283	4	0.14	92	290	4	0.09	89	279	4	0.12
voter	1838	8596	13	2.23	1784	8624	13	4.14	1838	8583	13	2.32	1777	8426	13	4.82
Improvement					2.57%	-2.57%	1.04%		-8.13%	-7.87%	7.52%		2.20%	-0.30%	12.39%	
Total				58.40				149.83				68.70				142.52

TABLE IV
LUT MAPPING IN THE EPFL SYNTHESIS COMPETITION.

Benchmark	Best [29]		dch -f; if -K 6		dch -f; if -Z 6 -K 10	
	LUTs	Depth	LUTs	Depth	LUTs	Depth
adder	347	5	360	6	445	5
bar	512	4	512	4	512	4
div	25318	175	23461	192	31526	175
hyp	182723	483	122394	511	154903	473
log2	8617	52	8778	60	9613	51
max	1114	6	1113	7	1250	6
multiplier	7785	25	6839	28	6903	25
sin	680530	10	1820	33	2379	27
sqrt	29593	162	30945	172	41626	156
square	3732	10	4189	11	4275	10

D. EPFL synthesis competition

This experiment shows that mapping using ACD can improve well-optimized LUT networks, resulting in best known results for 4 benchmarks in the EPFL synthesis competition. The previous best results were obtained using a portfolio of heavy logic optimizations applied to various representations, such as AIGs and LUT networks. In recent years, results have been further improved using *design-space exploration* (DSE) techniques that incrementally generate optimization scripts.

We obtain the optimized AIGs by repeatedly running the script used in the baseline of Table III along with additional delay-oriented AIG commands in ABC. From the obtained AIG, we compare traditional LUT mapping with choices to LUT mapping with ACD. Notably, results from the traditional mapper are quite far from the best results. This observation shows, as expected, that our technology-independent optimization finds worse AIGs than those used to obtain the best results. However, LUT mapping with ACD matches or

improves the depth for almost all benchmarks. The improved benchmarks are *hyp*, *log2*, *multiplier*, and *square*. Remarkably, our method reduces the depth of *hyp* by 10 levels, compared to the state of the art, while reducing area by 15%. In the benchmark *multiplier*, our result matches the depth but improves the number of LUTs. Benchmark *sin* is the only one where there is a large gap compared to the best result. In particular, the best result for *sin* requires significant logic duplication that is not performed in our synthesis flow. Contrarily to many other methods used to produce the best results, our results in Table III are obtained directly by LUT mapping without employing post-mapping optimization.

VI. CONCLUSION

This work proposes a novel formulation of Ashenhurst-Curtis decomposition (ACD) that enables efficient technology mapping and post-mapping resynthesis. The algorithm is truth-table-based and works for any size of the free set, bound set, and shared set, which makes it well-suited for delay optimization. We have shown that the proposed Boolean decomposition improves state-of-the-art in the decomposition quality with a competitive runtime. We have implemented and integrated the proposed method into a delay-driven LUT mapper. The experiments have shown that LUT mapping with ACD can improve the average delay by 12.39%, compared to the traditional structural LUT mapping with choices. Furthermore, the proposed approach has produced best results for 4 test cases in the EPFL synthesis competition.

ACKNOWLEDGMENTS

This research was supported by the SNF grant “Supercool: Design methods and tools for superconducting electronics”, 200021_1920981, and Synopsys Inc.

REFERENCES

- [1] R. L. Ashenhurst, "The decomposition of switching functions," 1957, pp. 74–116.
- [2] J. P. Curtis, "A new approach to the design of switching circuits," 1962.
- [3] J. P. Roth and R. M. Karp, "Minimization over boolean graphs," *IBM Journal of Research and Development*, vol. 6, no. 2, pp. 227–238, 1962.
- [4] V. N. Kravets and K. A. Sakallah, "Constructive multi-level synthesis by way of functional properties," Ph.D. dissertation, 2001.
- [5] C. Legl, B. Wurth, and K. Eckl, "Computing support-minimal subfunctions during functional decomposition," *Trans. VLSI*, vol. 6, no. 3, pp. 354–363, 1998.
- [6] M. Perkowski, M. Marek-Sadowska, L. Jozwiak, T. Luba, S. Grygiel, M. Nowicka, R. Malvi, Z. Wang, and J. Zhang, "Decomposition of multiple-valued relations," in *Proc. Inter. Symp. on Mult.- Valued Logic*, 1997, pp. 13–18.
- [7] J.-H. Jiang, Y. Jiang, and R. K. Brayton, "An implicit method for multi-valued network encoding," in *Proc. IWLS*, 2001, pp. 127–131.
- [8] R. Bryant, "Graph-based algorithms for boolean function manipulation," *IEEE Trans. on Computers*, vol. C-35, no. 8, pp. 677–691, 1986.
- [9] A. Mishchenko, R. Brayton, and S. Chatterjee, "Boolean factoring and decomposition of logic networks," in *Proc. ICCAD*, 2008, pp. 38–44.
- [10] S. Ray, A. Mishchenko, N. Een, R. Brayton, S. Jang, and C. Chen, "Mapping into LUT structures," in *Proc. DATE*, 2012.
- [11] J. Cong and Y. Ding, "FlowMap: an optimal technology mapping algorithm for delay optimization in lookup-table based FPGA designs," *Trans. CAD*, vol. 13, no. 1, pp. 1–12, 1994.
- [12] A. H. Farrahi and M. Sarrafzadeh, "Complexity of the lookup-table minimization problem for FPGA technology mapping," *IEEE Trans. CAD*, 1994.
- [13] E. Lehman, Y. Watanabe, J. Grodstein, and H. Harkness, "Logic decomposition during technology mapping," *Trans. CAD*, 1997.
- [14] G. Chen and J. Cong, "Simultaneous logic decomposition with technology mapping in FPGA designs," in *Proc. FPGA*, 2001, p. 48–55.
- [15] S. Chatterjee, A. Mishchenko, R. Brayton, X. Wang, and T. Kam, "Reducing structural bias in technology mapping," in *Proc. ICCAD*, 2005.
- [16] A. Mishchenko and R. Brayton, "Scalable logic synthesis using a simple circuit structure," in *Proc. IWLS*, 2006.
- [17] J. Cong, C. Wu, and Y. Ding, "Cut ranking and pruning: Enabling a general and efficient FPGA mapping solution," in *Proc. FPGA*, 1999.
- [18] A. Mishchenko, S. Cho, S. Chatterjee, and R. Brayton, "Combinational and sequential mapping with priority cuts," in *Proc. ICCAD*, 2007.
- [19] G. De Micheli, R. Brayton, and A. Sangiovanni-Vincentelli, "Optimal state assignment for finite state machines," *Trans. CAD*, vol. 4, no. 3, pp. 269–285, 1985.
- [20] T. Villa and A. Sangiovanni-Vincentelli, "NOVA: state assignment of finite state machines for optimal two-level logic implementation," *Trans. CAD*, vol. 9, no. 9, pp. 905–924, 1990.
- [21] S. Yang and M. Ciesielski, "Optimum and suboptimum algorithms for input encoding and its relationship to logic minimization," *Trans. CAD*, vol. 10, no. 1, pp. 4–12, 1991.
- [22] R. Brayton and A. Mishchenko, "ABC: An academic industrial-strength verification tool," in *Computer Aided Verification*, T. Touili, B. Cook, and P. Jackson, Eds., 2010. [Online]. Available: <https://github.com/berkeley-abc/abc>
- [23] L. Amarù, P.-E. Gaillardon, and G. D. Micheli, "The EPFL combinational benchmark suite," in *Proc. IWLS*, 2015.
- [24] A. Mishchenko, S. Chatterjee, and R. Brayton, "FRAIGs: A unifying representation for logic synthesis and verification," EECS Dep., UC Berkeley, Tech. Rep., 2005.
- [25] W. Yang, L. Wang, and A. Mishchenko, "Lazy man's logic synthesis," in *Proc. ICCAD*, 2012, p. 597–604.
- [26] Z. Huang, L. Wang, Y. Nasikovskiy, and A. Mishchenko, "Fast boolean matching based on NPN classification," in *Intern. Conf. on Field-Programmable Technology*, 2013.
- [27] M. Soeken, A. Mishchenko, A. Petkovska, B. Sterin, P. Jenne, R. K. Brayton, and G. De Micheli, "Heuristic NPN classification for large functions using AIGs and LEXSAT," in *Theory and Applications of Satisfiability Testing*, N. Creignou and D. Le Berre, Eds., 2016.
- [28] L. Fan and C. Wu, "FPGA technology mapping with adaptive gate decomposition," in *Proc. FPGA*, 2023, p. 135–140.
- [29] "EPFL synthesis competition best results [2023]." [Online]. Available: https://github.com/lisil/benchmarks/tree/v2023.1/best_results
- [30] A. Mishchenko, R. Brayton, J.-H. R. Jiang, and S. Jang, "Scalable don't-care-based logic optimization and resynthesis," *ACM Trans. Reconfigurable Technol. Syst.*, vol. 4, no. 4, 2011.
- [31] B. Schmitt, A. Mishchenko, and R. Brayton, "SAT-based area recovery in structural technology mapping," in *Proc. ASP-DAC*, 2018, pp. 586–591.