On the Implementation of Boolean Functions on Content-Addressable Memories

Ron M. Roth

***Abstract*—Let $[q\rangle$ denote the integer set $\{0, 1, \ldots, q-1\}$ and let $\mathbb{B} = \{0, 1\}$. The problem of implementing functions $[q\rangle \to \mathbb{B}$ on content-addressable memories (CAMs) is considered. CAMs can be classified by the input alphabet and the state alphabet of their cells; for example, in binary CAMs, those alphabets are both $\mathbb{B}$, while in a ternary CAM (TCAM), both alphabets are endowed with a "don't care" symbol.**

This work is motivated by recent proposals for using CAMs for fast inference on decision trees. In such learning models, the tree nodes carry out integer comparisons, such as testing equality ($x = t$?) or inequality ($x \le t$?), where $x \in [q\rangle$ is an input to the node and $t \in [q\rangle$ is a node parameter. A CAM implementation of such comparisons includes mapping (i.e., encoding) t into internal states of some number n of cells and mapping x into inputs to these cells, with the goal of minimizing n.

Such mappings are presented for various comparison families, as well as for the set of all functions $[q\rangle \to \mathbb{B}$, under several scenarios of input and state alphabets of the CAM cells. All those mappings are shown to be optimal in that they attain the smallest possible n for any given q.

***Index Terms*—Content-addressable memories, Integer comparisons, Representation of Boolean functions, VC dimension.**

I. Introduction

For $a, b \in \mathbb{Z}$, denote by $[a : b]$ the integer subset $\{z \in \mathbb{Z} : a \le z \le b\}$ and by $[a : b\rangle$ the set $[a : b-1]$; we will use the shorthand notation $[b\rangle$ for $[0 : b\rangle$.

Let $\mathbb{B}$ denote the Boolean set $\{0, 1\}$ and let $\mathbb{B}_*$ and $\mathbb{B}_\bullet$ denote the alphabets $\mathbb{B} \cup \{*\}$ and $\mathbb{B} \cup \{*, \bullet\}$, respectively. We refer to $*$ as the "don't-care" symbol and to $\bullet$ as the "reject" symbol. Define the function $\mathrm{T} : \mathbb{B}_\bullet^2 \to \mathbb{B}$ by the truth table shown in Table I. The first argument, u, of T is called the *input* and the second argument, ϑ, is the *internal state*. While the function $(u, \vartheta) \mapsto \mathrm{T}(u, \vartheta)$ is symmetric in its two arguments (in that $\mathrm{T}(u, \vartheta) = \mathrm{T}(\vartheta, u)$), there is a distinction between them in the way this function will be used.

TABLE I
Truth table of the function $(u, \vartheta) \mapsto \mathrm{T}(u, \vartheta)$.

$u \backslash \vartheta$	$\bullet$	0	1	$*$
$\bullet$	0	0	0	1
0	0	1	0	1
1	0	0	1	1
$*$	1	1	1	1

Ron M. Roth is with the Computer Science Department, Technion, Haifa 3200003, Israel. This work was done in part while visiting Hewlett Packard Laboratories, Milpitas, CA. This work was supported in part by Grant 1713/20 from the Israel Science Foundation. Part of this work was accepted for presentation at the IEEE Int'l Symposium on Information Theory (ISIT), June 2023. Email: ronny@cs.technion.ac.il.

We extend the definition of the function T to words (vectors) over $\mathbb{B}_\bullet$ as follows: for $\boldsymbol{u} = (u_j)_{j \in [n\rangle}$ and $\boldsymbol{\vartheta} = (\vartheta_j)_{j \in [n\rangle}$ in $\mathbb{B}_\bullet^n$, we define $\mathrm{T}(\boldsymbol{u}, \boldsymbol{\vartheta})$ by

$$\mathrm{T}(\boldsymbol{u}, \boldsymbol{\vartheta}) = \bigwedge_{j \in [n\rangle} \mathrm{T}(u_j, \vartheta_j), \tag{1}$$

where $\wedge$ denotes Boolean conjunction (i.e., logical "and").

A *content-addressable memory* (CAM) is a device which consists of an $m \times \ell$ array of *cells*, with each cell $(i, j) \in [m\rangle \times [\ell\rangle$ implementing the function $u \mapsto \mathrm{T}(u, \vartheta_{i,j})$, for some internal state value $\vartheta_{i,j} \in \mathbb{B}_\bullet$. The input to the array is a word $\boldsymbol{u} = (u_j)_{j \in [\ell\rangle} \in \mathbb{B}_\bullet^\ell$, with u_j serving as the input to all the cells along column j. The internal states need to be *programmed* into the array prior to its use and are therefore assumed to change much less frequently than the input word $\boldsymbol{u}$. In addition, the selections of the internal state values and of the input words are assumed to be independent.

Each row ("match line") in the array computes the conjunction of the outputs of the cells along the row (as in (1)), and the results obtained by the m rows form the output word, $\boldsymbol{v} \in \mathbb{B}^m$, of the array. The nonzero entries in $\boldsymbol{v}$ are referred to as "matches," and in practice there are settings where one might be interested only in the number of matches (namely, the Hamming weight of $\boldsymbol{v}$) or in the index i of the first match in $\boldsymbol{v}$.

CAMs are classified, *inter alia*, by the ranges of the input values u_j and the internal state values $\vartheta_{i,j}$. In a binary CAM (in short, BCAM), these values are constrained to be in $\mathbb{B}$, and in a ternary CAM (in short, TCAM), they are elements of $\mathbb{B}_*$. BCAMs are used in implementing fast look-up tables, and TCAMs are used in networking equipment for performing high-speed packet classification.

There are several common hardware designs for CAM cells, e.g., [6],[7],[13]. Most designs allow in effect the cells (including BCAM cells) to take both $*$ and $\bullet$ as input values. Moreover, some of the TCAM designs, such as the one shown in [6, Fig. 2(a)] and [7, Fig. 1], allow the internal state to take the value $\bullet$ (hence our election to define T to have the domain $\mathbb{B}_\bullet^2$).

For an integer $q \ge 2$, let $\mathcal{F}_q$ denote the set of all functions $f : [q\rangle \to \mathbb{B}$ (to avoid trivialities, we exclude the case $q = 1$).[1] For certain function families $\Phi \subseteq \mathcal{F}_q$, we are interested in the implementation of each $x \mapsto f(x) \in \Phi$ in the form

$$f(x) = \mathrm{T}(\boldsymbol{u}(x), \boldsymbol{\vartheta}(f)) \stackrel{(1)}{=} \bigwedge_{j \in [n\rangle} \mathrm{T}(u_j(x), \vartheta_j(f)), \tag{2}$$

[1] When $q = 2^h$, one can regard $\mathcal{F}_q$ as the set of all h-variate Boolean functions $\mathbb{B}^h \to \mathbb{B}$ (by simply replacing each element in the domain $[q\rangle$ by its length-h binary representation).

where $\boldsymbol{u} : [q\rangle \to \mathbb{B}_\bullet^n$ and $\boldsymbol{\vartheta} : \Phi \to \mathbb{B}_\bullet^n$ are prescribed mappings which may (only) depend on Φ and n; in fact, for a given family Φ, we will be interested in the smallest possible n for which any function $f \in \Phi$ can be expressed in the form (2). Note that the first argument of $\mathrm{T}(\cdot,\cdot)$ in (2) depends only on x, while the second depends only on f. Such a separation is dictated by the common use of CAMs: the functions f are pre-specified and the respective words $\boldsymbol{\vartheta}(f)$ are programmed into (the internal state values of) n CAM cells along a row, while x is an input within a sequence of inputs to which the functions are to be applied.

Our study of implementations of Boolean functions on CAMs is motivated primarily by certain models of machine learning—specifically, by recent proposals for using CAMs for fast inference on decision trees [8],[12]. In such learning models, the basic operations carried out in the nodes of the trees are comparisons between (real or) integer numbers. The implementation in [8],[12] can be seen as a generalized notion of a CAM which consists of an $m \times \ell$ array A of cells, with each cell $(i, s) \in [m\rangle \times [\ell\rangle$ computing a function (comparator) $x \mapsto f_{i,s}(x) \in \mathcal{F}_q$ which tests x against some fixed threshold. The input to such an array is a word $\boldsymbol{x} = (x_s)_{s\in[\ell\rangle} \in [q\rangle^\ell$ and the output of each row i is the conjunction of the outputs of the cells along the row:

$$\bigwedge_{s\in[\ell\rangle} f_{i,s}(x_s). \tag{3}$$

When used to implement decision trees, the entries of $\boldsymbol{x}$ are the features (attributes) and each row i corresponds to (a path that leads to) a leaf in the tree. The expression (3) evaluates to 1 if and only if the feature values in $\boldsymbol{x}$ lead to the (unique) leaf that corresponds to row i.

Our families Φ of interest have been selected accordingly and, in order to define them, we introduce the following notation. Let the bivariate functions E_q, N_q, Γ_q, and L_q, all with the domain $[q\rangle^2$ and range $\mathbb{B}$, be defined for every $(x, t) \in [q\rangle^2$ by:

$$\begin{aligned} \mathrm{E}_q(x,t) &= [\![x = t]\!], \\ \mathrm{N}_q(x,t) &= [\![x \neq t]\!], \\ \Gamma_q(x,t) &= [\![x \geq t]\!], \\ \mathrm{L}_q(x,t) &= [\![x \leq t]\!], \end{aligned}$$

where $[\![\cdot]\!]$ denotes the Iverson bracket (which evaluates to 1 if its argument is true, and to 0 otherwise). Respectively, define the following subsets of $\mathcal{F}_q$ (each of size q):

$$\begin{aligned} \mathcal{E}_q &= \left\{x \mapsto \mathrm{E}_q(x,t)\right\}_{t\in[q\rangle}, \\ \mathcal{N}_q &= \left\{x \mapsto \mathrm{N}_q(x,t)\right\}_{t\in[q\rangle}, \\ \mathcal{G}_q &= \left\{x \mapsto \Gamma_q(x,t)\right\}_{t\in[q\rangle}, \\ \mathcal{L}_q &= \left\{x \mapsto \mathrm{L}_q(x,t)\right\}_{t\in[q\rangle}. \end{aligned}$$

The families Φ that we consider are:

$$\mathcal{E}_q, \ \mathcal{N}_q, \ \mathcal{G}_q \ (\text{or } \mathcal{L}_q), \ \mathcal{G}_q \cup \mathcal{L}_q,$$

as well as the whole set $\mathcal{F}_q$.[2]

[2] The family $\mathcal{G}_q \cup \mathcal{L}_q$ is quite prevailing in decision trees, where the same feature x_s is typically compared against both lower and upper thresholds.

Once the functions in the selected family Φ have an implementation of the form (2) (with mappings $x \mapsto \boldsymbol{u}(x)$ and $f \mapsto \boldsymbol{\vartheta}(f)$ that attain the smallest possible n), we can realize an $m \times \ell$ array A of functions $f_{i,s} \in \Phi$ by an (ordinary) CAM consisting of an $m \times \ell n$ array $\tilde{\mathsf{A}}$ of cells, with each cell implementing the function $u \mapsto \mathrm{T}(u, \vartheta_{i,j})$ with the internal states $\vartheta_{i,j}$ selected so that non-overlapping n-cell blocks in $\tilde{\mathsf{A}}$ implement the functions $f_{i,s}$.[3] Specifically, we select the internal states in $\tilde{\mathsf{A}}$ for each $(i, j) \in [m\rangle \times [\ell n\rangle$ so that

$$\left(\vartheta_{i,sn} \ \vartheta_{i,sn+1} \ \ldots \ \vartheta_{i,(s+1)n-1}\right) = \boldsymbol{\vartheta}(f_{i,s}),$$

in which case the output of row i in A (namely, (3)) can be expressed as

$$\begin{aligned} \bigwedge_{s\in[\ell\rangle} f_{i,s}(x_s) &\overset{(2)}{=} \bigwedge_{s\in[\ell\rangle} \mathrm{T}(\boldsymbol{u}(x_s), \boldsymbol{\vartheta}(f_{i,s})) \\ &= \bigwedge_{s\in[\ell\rangle} \bigwedge_{j\in[n\rangle} \mathrm{T}(u_j(x_s), \vartheta_{i,sn+j}), \end{aligned}$$

which is the output of row i in $\tilde{\mathsf{A}}$.

Based upon the intended use of CAMs and the various designs of their cells, we consider several scenarios, which will be identified by pairs of symbols from $\{\circ, *, \bullet\}$: the first (respectively, second) symbol in such a pair specifies the alphabet of the range of the mapping $\boldsymbol{u}$ (respectively, $\boldsymbol{\vartheta}$) that appears in (2), with $\mathbb{B}_\circ = \mathbb{B}$; for example, in Scenario $(\circ *)$, the ranges of $\boldsymbol{u}$ and $\boldsymbol{\vartheta}$ are $\mathbb{B}^n$ and $\mathbb{B}_*^n$, respectively. We will then say that a subset $\Phi \subseteq \mathcal{F}_q$ is *n-cell implementable* under a given scenario if there exist mappings $\boldsymbol{u}$ and $\boldsymbol{\vartheta}$ with the appropriate ranges ($\mathbb{B}^n$, $\mathbb{B}_*^n$, or $\mathbb{B}_\bullet^n$) such that every function $f \in \Phi$ can be written in the form (2).

A. Summary of results

Table II summarizes our results: for each of our selected choices for Φ and for six possible scenarios, the table shows the largest value of q for which Φ is n-cell implementable. Due to the symmetries $\mathrm{E}_q(x,t) = \mathrm{E}_q(t,x)$, $\mathrm{N}_q(x,t) = \mathrm{N}_q(t,x)$, and $\Gamma_q(x,t) = \Gamma_q(q-1-t, q-1-x) = \mathrm{L}_q(t,x)$, the columns $(*\circ)$ and $(\circ *)$ are the same in the rows that correspond to the families $\mathcal{E}_q$, $\mathcal{N}_q$, and $\mathcal{G}_q$, and so are the columns $(\bullet *)$ and $(*\bullet)$. Hence, for these families, there are four scenarios to consider.

Three scenarios were omitted from the table: Scenario $(\circ\circ)$ corresponds to a BCAM, where the two arguments of $\mathrm{T}(\cdot,\cdot)$ are restricted to $\mathbb{B}$. Such a restriction limits the subsets Φ that can be expressed in the form (2) (regardless of how large n is), namely, if $f(y) = f(z) = 1$ for some $f \in \Phi$ and some two distinct $y, z \in [q\rangle$, then $g(y) = g(z)$ for all $g \in \Phi$. Among our families of interest, only $\mathcal{E}_q$ satisfies this property (see Footnote 4 below). Scenarios $(\circ\bullet)$ and $(\bullet\circ)$ are also omitted.

[3] The particular architecture of CAMs that is proposed [8],[12], referred to therein as an analog CAM (a-CAM), consists of cells which perform comparisons between integers in $[q\rangle$, for some prescribed q. The design of the cells makes use of programmable resistors and the comparisons are carried out in the analog domain. The reliability of such an a-CAM in practice, however, is yet to be tested and, in any event, such an a-CAM is expected to operate only for relatively small values of q. The alternate approach, which motivates this work, proposes using more traditional CAM designs (which are well tested and robust) to realize the same (targeted) functionality of an a-CAM.

Scenario (∘•) is identical to (∘∗) unless Φ contains the all-zero function, and among our families of interest, this occurs only when $\Phi = \mathcal{F}_q$; in this case Scenarios (∘•) and (∘∗) differ only in one isolated case (see Footnote 8 below). Similarly, Scenarios (•∘) and (∗∘) would differ only if there were an element $y \in [q\rangle$ such that $f(y) = 0$ for all $f \in \Phi$; yet this does not happen in any of the families Φ of interest.

We see from the table that for a given $q \geq 3$ and under Scenarios (∗∘), (∘∗), (∗•), and (••), the family $\mathcal{N}_q$ is n-cell implementable, (if and) only if the *whole* set $\mathcal{F}_q$ is. As such, the family $\mathcal{N}_q$, albeit forming an exponentially small fraction of $\mathcal{F}_q$ (q out of the 2^q functions), necessarily has the least-efficient CAM implementation under these scenarios. And this "record" is almost tied by the family $\mathcal{G}_q$ (which, too, is of size q). On the other hand, under Scenarios (∗∗) and (•∗), each of the families $\mathcal{N}_q$, $\mathcal{G}_q$ (or even $\mathcal{G}_q \cup \mathcal{L}_q$) is $\lceil q/2 \rceil$-cell implementable, while the whole set $\mathcal{F}_q$ requires (at least) q cells (i.e., for this set, Scenarios (∗∗) and (•∗) do not offer any advantage over Scenarios (∗∘) or (∘∗)).

The problem of constructing efficient encodings of integer intervals on a TCAM (which include in particular efficient implementations of integer comparisons) has been the studied in the literature, e.g., [4],[5]. However, in those papers, the encoding of an interval indicator (or of a comparison) can utilize several rows of the TCAM. This translates into allowing taking a disjunction ("or", "$\vee$") of several terms of the form seen in the right-hand side of (2), with each term having its own word $\boldsymbol{\vartheta}$ (the common mapping $x \mapsto \boldsymbol{u}(x)$ is usually taken as the ordinary binary representation of x). However, in our setting, each TCAM row may contain (the conjunction of) the implementation of several comparators and, therefore, any such implementation cannot be split across several rows.

Our families of interest, $\mathcal{E}_q$, $\mathcal{N}_q$, $\mathcal{G}_q$, $\mathcal{G}_q \cup \mathcal{L}_q$, and $\mathcal{F}_q$ are covered, respectively, in Sections II through VI (followed by a discussion in Section VII). In each section except Section VI, the scenarios are covered in the order they appear in Table II (skipping Scenarios (∗∘) and (•∗) when they coincide with Scenarios (∘∗) and (∗•)). Some effort has been put into making the text less tedious, e.g., by combining several proofs. We substantiate the values in Table II by presenting both lower and upper bounds for each entry. The lower bounds are all obtained by presenting explicit mappings $x \mapsto \boldsymbol{u}(x)$ and $f \mapsto \boldsymbol{\vartheta}(f)$ for which Eq. (2) holds for the respective family Φ. Admittedly, many of the optimal mappings are rather straightforward (e.g., they are either standard binary representations of their argument or kinds of unary representations, as in the proofs of Propositions 1–3 below); in those cases, the contribution of this work lies mainly in showing that these mappings are, in fact, the best possible. Notable exceptions are the mappings for Scenarios (•∗), (∗•), and (••) when $\Phi = \mathcal{E}_q$, and for Scenarios (∗∗) and (•∗) when $\Phi = \mathcal{G}_q \cup \mathcal{L}_q$ (and, perhaps, to some extent, Scenario (∗∗) when $\Phi = \mathcal{G}_q$). As part of our study of the whole set $\mathcal{F}_q$ in Section VI, we also draw a connection between the subject of this paper and that of computing the VC dimension of Boolean monomials.

B. Some simple observations

We end our introduction by stating three simple observations which will be useful in the sequel.

Proposition 1. *Under each of the scenarios in Table II, any subset $\Phi \subseteq \mathcal{F}_q$ is n-cell implementable, whenever*

$$q \leq n.$$

Proof. Given that $q \leq n$, we can extend (arbitrarily) the functions in $\mathcal{F}_q$ so that they are defined over the (possibly larger) domain $[n\rangle$ and, so, we assume that $q = n$. For Scenario (∘∗), we take the mappings $x \mapsto \boldsymbol{u}(x) = (u_j(x))_{j\in[n\rangle}$ and $f \mapsto \boldsymbol{\vartheta}(f) = (\vartheta_j(f))_{j\in[n\rangle}$ to be

$$u_j(x) = \begin{cases} 1 & \text{if } x = j \\ 0 & \text{if } x \neq j \end{cases} \quad \text{and} \quad \vartheta_j(f) = \begin{cases} * & \text{if } f(j) = 1 \\ 0 & \text{if } f(j) = 0 \end{cases}.$$

Otherwise, we take them to be

$$u_j(x) = \begin{cases} 1 & \text{if } x = j \\ * & \text{if } x \neq j \end{cases} \quad \text{and} \quad \vartheta_j(f) = f(j).$$

□

Taking the family $\Phi = \mathcal{N}_q = \{\mathrm{N}_q(\cdot, t)\}_{t\in[q\rangle}$ as an example, Table III shows the respective mappings $x \mapsto \boldsymbol{u}(x)$ and $\mathrm{N}_q(\cdot, t) \mapsto \boldsymbol{\vartheta}(\mathrm{N}_q(\cdot, t))$ under Scenario (∘∗).

TABLE III
MAPPINGS FOR $\Phi = \mathcal{N}_q$ THAT ATTAIN $q = n$ UNDER SCENARIO (∘∗).

x	$\boldsymbol{u}(x)$
0	$10000\ldots0$
1	$01000\ldots0$
2	$00100\ldots0$
$\vdots$	$\ddots$
$n-2$	$000\ldots010$
$n-1$	$0000\ldots01$

t	$\boldsymbol{\vartheta}(\mathrm{N}_q(\cdot, t))$
0	$0{*}{*}{*}{*}\ldots{*}$
1	${*}0{*}{*}{*}\ldots{*}$
2	${*}{*}0{*}{*}\ldots{*}$
$\vdots$	$\ddots$
$n-2$	${*}{*}{*}\ldots{*}0{*}$
$n-1$	${*}{*}{*}{*}\ldots{*}0$

Proposition 2. *Under each of the scenarios in Table II, any subset $\Phi \subseteq \mathcal{F}_q$ is n-cell implementable, whenever*

$$|\Phi| \leq n.$$

Proof. Again, we can assume that $|\Phi| = n$. Writing $\Phi = \{f_1, f_2, \ldots, f_{|\Phi|}\}$, for Scenario (∗∘), we take the mappings $x \mapsto \boldsymbol{u}(x)$ and $f \mapsto \boldsymbol{\vartheta}(f)$ to be

$$u_j(x) = \begin{cases} * & \text{if } f_j(x) = 1 \\ 0 & \text{if } f_j(x) = 0 \end{cases} \quad \text{and} \quad \vartheta_j(f) = \begin{cases} 1 & \text{if } f = f_j \\ 0 & \text{if } f \neq f_j \end{cases}.$$

Otherwise, we take them to be

$$u_j(x) = f_j(x) \quad \text{and} \quad \vartheta_j(f) = \begin{cases} 1 & \text{if } f = f_j \\ * & \text{if } f \neq f_j \end{cases}.$$

□

Proposition 3. *Under each of the scenarios (∗•) or (••), any subset $\Phi \subseteq \mathcal{F}_q$ is n-cell implementable, whenever*

$$q \leq 2n.$$

TABLE II
LARGEST VALUE OF q THAT CAN BE ACCOMMODATED FOR ANY NUMBER OF CELLS n.

<table>
<tr><th>Φ</th><th>Scenario $(*\circ)$</th><th>Scenario $(\circ *)$</th><th>Scenario $(**)$</th><th>Scenario $(\bullet *)$</th><th>Scenario $(*\bullet)$</th><th>Scenario $(\bullet\bullet)$</th></tr>
<tr><td>$\mathcal{E}_q$</td><td colspan="3">2^n</td><td colspan="2">$\binom{n}{\lfloor n/3 \rfloor} \cdot 2^{\lceil 2n/3 \rceil}$</td><td>$\binom{2n}{n}$</td></tr>
<tr><td>$\mathcal{N}_q$</td><td colspan="2">$2 \quad (n=1)$
$n \quad (n \geq 2)$</td><td colspan="4">$2n$</td></tr>
<tr><td>$\mathcal{G}_q$</td><td colspan="2">$n+1$</td><td>$2n \quad (n=1,2)$
$2n{+}1 \quad (n \geq 3)$</td><td colspan="3">$2n+1$</td></tr>
<tr><td>$\mathcal{G}_q \cup \mathcal{L}_q$</td><td>$n$</td><td>$n+1$</td><td colspan="2">$2 \quad (n=1)$
$2n-1 \quad (n \geq 2)$</td><td colspan="2" rowspan="2">$2n$</td></tr>
<tr><td>$\mathcal{F}_q$</td><td colspan="4">n</td></tr>
</table>

Proof. Assuming that $q = 2n$, we take the mappings $x \mapsto \boldsymbol{u}(x)$ and $f \mapsto \boldsymbol{\vartheta}(f)$ to be

$$u_j(x) = \begin{cases} 0 & \text{if } x = 2j \\ 1 & \text{if } x = 2j+1 \\ * & \text{otherwise} \end{cases} \tag{4}$$

and

$$\vartheta_j(f) = \begin{cases} * & \text{if } f(2j) = f(2j{+}1) = 1 \\ 0 & \text{if } f(2j) = 1 \text{ and } f(2j{+}1) = 0 \\ 1 & \text{if } f(2j) = 0 \text{ and } f(2j{+}1) = 1 \\ \bullet & \text{if } f(2j) = f(2j{+}1) = 0 \end{cases} . \tag{5}$$

□

II. THE FAMILY $\mathcal{E}_q$

In this section, we consider the family $\Phi = \mathcal{E}_q$. Since each function $x \mapsto \mathrm{E}_q(x,t) \in \mathcal{E}_q$ can be identified by the parameter $t \in [q\rangle$, we will use for convenience the notation $t \mapsto \boldsymbol{\vartheta}(t)$ for $\mathrm{E}_q(\cdot,t) \mapsto \boldsymbol{\vartheta}(\mathrm{E}_q(\cdot,t))$. Due to the symmetry $\mathrm{E}_q(x,t) = \mathrm{E}_q(t,x)$, it will suffice to state our results only for Scenarios $(\circ *)$, $(**)$, $(*\bullet)$, and $(\bullet\bullet)$.

The next proposition handles the case $\Phi = \mathcal{E}_q$ under Scenarios $(\circ *)$ and $(**)$.

Proposition 4. *Under each of the scenarios $(\circ *)$ or $(**)$, the family $\mathcal{E}_q$ is n-cell implementable, if and only if*

$$q \leq 2^n.$$

Proof. Sufficiency follows by letting $\boldsymbol{u}(x)$ and $\boldsymbol{\vartheta}(t)$ be the length-n binary representations of x and t, respectively.[4]

Necessity under Scenario $(\circ *)$ follows from the fact that the mapping $x \mapsto \boldsymbol{u}(x)$ must be injective (otherwise, if $\boldsymbol{u}(x_0) = \boldsymbol{u}(x_1)$ for $x_0 \neq x_1$, then Eq. (2) would imply $\mathrm{E}_q(x_0,x_1) = \mathrm{E}_q(x_1,x_1) = 1$, which is a contradiction).

As for Scenario $(**)$, suppose that Eq. (2) holds for some $x \mapsto \boldsymbol{u}(x)$ and $t \mapsto \boldsymbol{\vartheta}(t)$. Define the mapping $x \mapsto \widehat{\boldsymbol{u}}(x) = (\widehat{u}_j(x))_{j \in [n\rangle}$ as follows:

$$\widehat{u}_j(x) = \begin{cases} u_j(x) & \text{if } u_j(x) \in \mathbb{B} \\ \vartheta_j(x) & \text{if } u_j(x) = * \text{ and } \vartheta_j(x) \in \mathbb{B} \\ 0 & \text{if } u_j(x) = \vartheta_j(x) = * \end{cases} .$$

It readily follows that $\mathrm{T}(\widehat{\boldsymbol{u}}(x), \boldsymbol{\vartheta}(x)) = \mathrm{T}(\boldsymbol{u}(x), \boldsymbol{\vartheta}(x)) = 1$ for every $x \in [q\rangle$. On the other hand, since $\widehat{\boldsymbol{u}}(x)$ is obtained from $\boldsymbol{u}(x)$ by (possibly) changing some entries from $*$ into elements of $\mathbb{B}$, then $\mathrm{T}(\boldsymbol{u}(x), \boldsymbol{\vartheta}(t)) = 0$ implies $\mathrm{T}(\widehat{\boldsymbol{u}}(x), \boldsymbol{\vartheta}(t)) = 0$. Hence, $\mathrm{E}_q(x,t) = \mathrm{T}(\widehat{\boldsymbol{u}}(x), \boldsymbol{\vartheta}(t))$ for all $x, t \in [q\rangle$, thereby reducing to Scenario $(\circ *)$. □

The family $\Phi = \mathcal{E}_q$ under Scenarios $(*\bullet)$ and $(\bullet\bullet)$ will be treated next, yet, to this end, we will need some definitions and two lemmas.

We introduce the following partial ordering, $\preceq$, on the elements of $\mathbb{B}_\bullet$:

$$\bullet \;\preceq\; 0,1 \;\preceq\; *$$

(with no ordering defined between 0 and 1), and extend it to words $\boldsymbol{v}, \boldsymbol{v}' \in \mathbb{B}_\bullet^n$, with $\boldsymbol{v} \preceq \boldsymbol{v}'$ if the relation holds componentwise.

A *(maximal) chain* over $\mathbb{B}_*^n$ is a list of $n+1$ words in $\mathbb{B}_*^n$,

$$(\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n),$$

where $\boldsymbol{v}_0$ is a word in $\mathbb{B}^n$ and for each $i \in [n\rangle$, the word $\boldsymbol{v}_{i+1}$ is obtained from $\boldsymbol{v}_i$ by changing one non-$*$ entry into a $*$; thus, $\boldsymbol{v}_n$ is always the all-$*$ word.[5] It is rather easy to see that the number of chains over $\mathbb{B}_*^n$ is $2^n \cdot n!$.

An *antichain* in $\mathbb{B}_*^n$ (respectively, $\mathbb{B}_\bullet^n$) is a subset $\mathcal{A}$ of $\mathbb{B}_*^n$ (respectively, $\mathbb{B}_\bullet^n$) such that $\boldsymbol{v} \not\preceq \boldsymbol{v}'$ for any two distinct words $\boldsymbol{v}, \boldsymbol{v}' \in \mathcal{A}$.

[4] Such mappings suit also Scenario $(\circ\circ)$.

[5] The definition of a chain can also be extended to $\mathbb{B}_\bullet^n$, but our analysis below will use a shortcut that will not require this extension.

For a word $\boldsymbol{v} \in \mathbb{B}_\bullet^n$ and a symbol $\sigma \in \mathbb{B}_\bullet$, we denote by $\mathsf{w}_\sigma(\boldsymbol{v})$ the number of occurrences of the symbol σ in $\boldsymbol{v}$. For $w \in [0:n]$, let

$$\mathbb{B}_*^n(w) = \left\{ \boldsymbol{v} \in \mathbb{B}_*^n \,:\, \mathsf{w}_0(\boldsymbol{v}) + \mathsf{w}_1(\boldsymbol{v}) = w \right\}.$$

The following lemma is a generalization of Sperner's theorem [1, p. 2].

Lemma 5. *Let $\mathcal{A}$ be an antichain in $\mathbb{B}_*^n$. Then*

$$|\mathcal{A}| \le \binom{n}{\lfloor n/3 \rfloor} \cdot 2^{\lceil 2n/3 \rceil},$$

with equality attained by the set $\mathbb{B}_*^n(\lceil 2n/3 \rceil)$.

Proof. We adapt Lubell's proof of Sperner's theorem (in [9]) to our setting. Let $\mathcal{A}$ be an antichain and for $w \in [0:n]$, let $\mathcal{A}(w) = \mathcal{A} \cap \mathbb{B}_*^n(w)$. We first observe that any word $\boldsymbol{v} \in \mathcal{A}(w)$ is contained in exactly

$$2^{n-w} \cdot w!(n-w)!$$

chains over $\mathbb{B}_*^n$ and none of the other words in those chains belongs to $\mathcal{A}$. Hence,

$$\sum_{w \in [0:n]} 2^{n-w} \cdot w!(n-w)! \cdot |\mathcal{A}(w)| \le 2^n \cdot n!,$$

which can also be written as

$$\sum_{w \in [0:n]} \frac{|\mathcal{A}(w)|}{\beta(w)} \le 1, \tag{6}$$

where

$$\beta(w) = 2^w \cdot \binom{n}{w}.$$

The latter expression is maximized when $w = w_{\max} = \lceil 2n/3 \rceil$ and, so,

$$\begin{aligned} |\mathcal{A}| &= \sum_{w \in [0:n]} |\mathcal{A}(w)| \\ &\le \sum_{w \in [0:n]} \frac{\beta(w_{\max})}{\beta(w)} \cdot |\mathcal{A}(w)| \\ &\overset{(6)}{\le} \beta(w_{\max}) = \binom{n}{\lfloor n/3 \rfloor} \cdot 2^{\lceil 2n/3 \rceil}. \end{aligned}$$

The inequalities become equalities when $\mathcal{A} = \mathbb{B}_*^n(\lceil 2n/3 \rceil)$ which is, indeed, an antichain. $\square$

The counterpart of Lemma 5 for $\mathbb{B}_\bullet^n$ takes the following form.

Lemma 6. *Let $\mathcal{A}$ be an antichain in $\mathbb{B}_\bullet^n$. Then*

$$|\mathcal{A}| \le \binom{2n}{n},$$

with equality attained by the set

$$\mathbb{B}_\circledast^n = \left\{ \boldsymbol{v} \in \mathbb{B}_\bullet^n \,:\, \mathsf{w}_\bullet(\boldsymbol{v}) = \mathsf{w}_*(\boldsymbol{v}) \right\}$$

(namely, $\mathbb{B}_\circledast^n$ is the set of all words in $\mathbb{B}_\bullet^n$ in which the symbols $\bullet$ and $*$ have the same count).

Proof. Consider the following mapping $\lambda : \mathbb{B}_\bullet \to \mathbb{B}^2$:

$$\lambda(\bullet) = 00, \quad \lambda(0) = 01, \quad \lambda(1) = 10, \quad \lambda(*) = 11.$$

We extend this definition to a mapping from words $\boldsymbol{v} = (v_j)_{j \in [n\rangle} \in \mathbb{B}_\bullet^n$ to words in $\mathbb{B}^{2n}$ by

$$\lambda(\boldsymbol{v}) = \lambda(v_0) \| \lambda(v_1) \| \ldots \| \lambda(v_{n-1})$$

(with $\|$ denoting concatenation) and, accordingly, from subsets of $\mathbb{B}_\bullet^n$ to subsets of $\mathbb{B}^{2n}$. Clearly, λ is a bijection under all these settings. Moreover, a subset $\mathcal{A} \subseteq \mathbb{B}_\bullet^n$ is an antichain in $\mathbb{B}_\bullet^n$, if and only if $\lambda(\mathcal{A})$ is an antichain in $\mathbb{B}^{2n}$ under the ordering $0 \le 1$ on the elements of $\mathbb{B}$; namely, there are no two distinct words $\boldsymbol{v}, \boldsymbol{v}' \in \mathcal{A}$ such that $\lambda(\boldsymbol{v}) \le \lambda(\boldsymbol{v}')$ (componentwise). Hence, by Sperner's theorem we have, for every antichain $\mathcal{A}$ in $\mathbb{B}_\bullet^n$,

$$|\mathcal{A}| = |\lambda(\mathcal{A})| \le \binom{2n}{n},$$

with equality holding when $\mathcal{A}$ is the following subset of $\mathbb{B}_\bullet^n$ (see [1, p. 2]):

$$\left\{ \boldsymbol{v} \in \mathbb{B}_\bullet^n \,:\, \mathsf{w}_0(\lambda(\boldsymbol{v})) = \mathsf{w}_1(\lambda(\boldsymbol{v})) = n \right\}.$$

One can easily verify that this set is identical to $\mathbb{B}_\circledast^n$. $\square$

We now turn to stating our result for the family $\Phi = \mathcal{E}_q$ under Scenarios $(*\bullet)$ and $(\bullet\bullet)$.

Proposition 7. *The following holds for the subset $\mathcal{E}_q$.*

(a) Under Scenario $(*\bullet)$, it is n-cell implementable, if and only if

$$q \le \binom{n}{\lfloor n/3 \rfloor} \cdot 2^{\lceil 2n/3 \rceil}.$$

(b) Under Scenario $(\bullet\bullet)$, it is n-cell implementable, if and only if

$$q \le \binom{2n}{n}.$$

Proof. Starting with proving necessity, we observe that the mapping $x \mapsto \boldsymbol{u}(x)$ has to be injective and its set of images has to be an antichain in $\mathbb{B}_*^n$ (for part (a)) or in $\mathbb{B}_\bullet^n$ (for part (b)): indeed, if we had $\boldsymbol{u}(x) \preceq \boldsymbol{u}(x')$ for distinct $x, x' \in [q\rangle$ then, from $\mathrm{E}_q(x,x) = \mathrm{T}(\boldsymbol{u}(x), \boldsymbol{\vartheta}(x)) = 1$ we would get the contradiction $\mathrm{E}_q(x',x) = \mathrm{T}(\boldsymbol{u}(x'), \boldsymbol{\vartheta}(x)) = 1$. The sought result then follows from Lemmas 5 and 6.

Sufficiency follows by selecting $x \mapsto \boldsymbol{u}(x)$ to be any injective mapping into $\mathbb{B}_*^n(\lceil 2n/3 \rceil)$ (for part (a)) or into $\mathbb{B}_\circledast^n$ (for part (b)) and defining $t \mapsto \boldsymbol{\vartheta}(t)$ for every $t \in [q\rangle$ as follows:

$$\vartheta_j(t) = \begin{cases} u_j(t) & \text{if } u_j(t) \in \mathbb{B} \\ \bullet & \text{if } u_j(t) = * \\ * & \text{if } u_j(t) = \bullet \end{cases}.$$

$\square$

Using known approximations for the binomial coefficients (see [10, p. 309]), the inequalities in parts (a) and (b) of Proposition 7 are implied by

$$q \le \frac{3}{4\sqrt{n}} \cdot 3^n \quad \text{and} \quad q \le \frac{1}{2\sqrt{n}} \cdot 4^n,$$

respectively. This means that q can almost get to the size of $\mathbb{B}_*^n$ (in part (a)) or of $\mathbb{B}_\bullet^n$ (in part (b)) (this size would have been reachable if $\mathrm{T}(\cdot,\cdot)$ had been defined so that $\mathrm{T}(u,\vartheta) = 1$ if and only if $u = \vartheta$, for any $u, \vartheta \in \mathbb{B}_\bullet$).

Remark 1. Referring to the mapping $\lambda : \mathbb{B}_\bullet \to \mathbb{B}^2$ in the proof of Lemma 6, for any $u \in \mathbb{B}_\bullet$, denote by $\lambda_0(u)$ and $\lambda_1(u)$ the first and second entries (in $\mathbb{B}$) of $\lambda(u)$. It is fairly easy to see that for every $u, \vartheta \in \mathbb{B}_\bullet$,

$$\mathrm{T}(u, \vartheta) = \big(\lambda_0(u) \vee \lambda_1(\vartheta)\big) \wedge \big(\lambda_1(u) \vee \lambda_0(\vartheta)\big).$$

In fact, this is (essentially) how TCAM cells are implemented in [6, Fig. 2(a)] and in [7, Fig. 1]. Given a family $\Phi \subseteq \mathcal{F}_q$, we can therefore rewrite (2) as

$$f(x) = \bigwedge_{j \in [2n\rangle} \left(\widetilde{u}_j(x) \vee \widetilde{\vartheta}_j(f)\right)$$

with mappings $\widetilde{\boldsymbol{u}} : [q\rangle \to \mathbb{B}^{2n}$ and $\widetilde{\boldsymbol{\vartheta}} : \Phi \to \mathbb{B}^{2n}$ that satisfy certain constraints, depending on the scenario (the only exception is Scenario ($\bullet\bullet$), where no constraints are imposed). $\square$

III. The family $\mathcal{N}_q$

In this section, we consider the family $\Phi = \mathcal{N}_q$. As we did in Section II, we will identify each function $x \mapsto \mathrm{N}_q(x,t) \in \mathcal{N}_q$ by the parameter $t \in [q\rangle$ and will use the notation $t \mapsto \boldsymbol{\vartheta}(t)$ for $\mathrm{N}_q(\cdot,t) \mapsto \boldsymbol{\vartheta}(\mathrm{N}_q(\cdot,t))$. And, here, too, it will suffice to state the results only for Scenarios ($\circ *$), ($**$), ($*\bullet$), and ($\bullet\bullet$).

Proposition 8. *Under Scenario* ($\circ *$)*, the family $\mathcal{N}_q$ is n-cell implementable, if and only if*

$$q \le \begin{cases} 2 & \text{if } n = 1 \\ n & \text{if } n \ge 2 \end{cases}.$$

Proof. The case $n = 1$ is easily verified, so we assume hereafter in the proof that $n \ge 2$.

Sufficiency follows from Proposition 1 or 2 (see Table III), and necessity follows essentially from the proof of Theorem 1 in [11]; we include a proof for completeness and for reference in the sequel. Suppose that Eq. (2) holds for $\mathcal{N}_q$ with mappings $\boldsymbol{u} : [q\rangle \to \mathbb{B}^n$ and $\boldsymbol{\vartheta} : \mathcal{N}_q \to \mathbb{B}_*^n$; both these mappings are necessarily injective. For every $t \in [q\rangle$ we have $\mathrm{T}(\boldsymbol{u}(t), \boldsymbol{\vartheta}(t)) = \mathrm{N}_q(t,t) = 0$ and, so, there exists an index $j \in [n\rangle$ such that $\vartheta_j(t) \in \mathbb{B}$ and $u_j(t) \ne \vartheta_j(t)$; moreover, there is no loss of generality in assuming that all the other entries in $\boldsymbol{\vartheta}(t)$ are $*$. Assume now to the contrary that $q > n \ge 2$. By the pigeonhole principle there exist distinct $t_0, t_1 \in [q\rangle$ for which that index j is the same. Since the mapping $\boldsymbol{\vartheta}$ is injective, we must have $\vartheta_j(t_0) \ne \vartheta_j(t_1)$. Yet then, for (any) $t_2 \ne t_0, t_1$ in $[q\rangle$ we have $u_j(t_2) \ne \vartheta_j(t_i)$ for some $i \in [2\rangle$, thereby yielding the contradiction $\mathrm{N}_q(t_2, t_i) = \mathrm{T}(\boldsymbol{u}(t_2), \boldsymbol{\vartheta}(t_i)) = 0$. $\square$

Recall from Propositions 1 and 2 that the whole set $\mathcal{F}_q$ is q-cell implementable and that any subset Φ of $\mathcal{F}_q$ of size $|\Phi| < q$ is $(q-1)$-cell implementable. Proposition 8 implies that for any $q \ge 3$ and under Scenarios ($*\circ$) and ($\circ *$), the family $\mathcal{N}_q$ is a smallest possible subset of $\mathcal{F}_q$ that requires the same number of cells, q, as the whole set $\mathcal{F}_q$ does.

Proposition 9. *Under each of the scenarios* ($**$), ($*\bullet$), *or* ($\bullet\bullet$)*, the family $\mathcal{N}_q$ is n-cell implementable, if and only if*

$$q \le 2n.$$

Proof. Sufficiency follows from Proposition 3 which, for the family $\mathcal{N}_q$, holds in fact also under Scenario ($**$): specifically, for this family, the range of the mapping $\boldsymbol{\vartheta}$ in (5) is actually $\mathbb{B}_*^n$.

We next prove necessity under (the loosest) Scenario ($\bullet\bullet$), and our arguments will be similar to those used in the previous proof. Specifically, assuming that Eq. (2) holds, for every $t \in [q\rangle$ there exists an index $j \in [n\rangle$ such that $u_j(t), \vartheta_j(t) \in \mathbb{B} \cup \{\bullet\}$ and either $\vartheta_j(t) = \bullet$ or $u_j(t) \ne \vartheta_j(t)$; moreover, all the other entries in $\boldsymbol{\vartheta}(t)$ can be assumed to be $*$. Suppose to the contrary that $q > 2n$. By the pigeonhole principle there exist distinct $t_0, t_1, t_2 \in [q\rangle$ for which that index j is the same. Since $\boldsymbol{\vartheta}$ is injective, we must have $\vartheta_j(t_i) = \bullet$ for (exactly) one $i \in [3\rangle$. Yet then $\mathrm{N}_q(t_k, t_i) = \mathrm{T}(\boldsymbol{u}(t_k), \boldsymbol{\vartheta}(t_i)) = 0$ for $k \in [3\rangle \setminus \{i\}$, which is a contradiction. $\square$

IV. The family $\mathcal{G}_q$

In this section, we consider the family $\Phi = \mathcal{G}_q$. We again identify each function $x \mapsto \Gamma_q(x,t) \in \mathcal{G}_q$ by the parameter $t \in [q\rangle$ and use the notation $t \mapsto \boldsymbol{\vartheta}(t)$ for $\Gamma_q(\cdot,t) \mapsto \boldsymbol{\vartheta}(\Gamma_q(\cdot,t))$. Clearly, $\Gamma_q(x,t) = \mathrm{T}(\boldsymbol{u}(x), \boldsymbol{\vartheta}(t))$ if and only if $\mathrm{L}_q(x,t) = \mathrm{T}(\boldsymbol{u}'(x), \boldsymbol{\vartheta}'(t))$, where $\boldsymbol{u}'(x) = \boldsymbol{u}(q{-}1{-}x)$ and $\boldsymbol{\vartheta}'(t) = \boldsymbol{\vartheta}(q{-}1{-}t)$; hence, our analysis will apply just as well to $\Phi = \mathcal{L}_q$. Moreover, since $\Gamma_q(x,t) = \mathrm{L}_q(t,x)$, it will suffice to state the results only for Scenarios ($\circ *$), ($**$), ($*\bullet$), and ($\bullet\bullet$).

Proposition 10. *Under Scenario* ($\circ *$)*, the family $\mathcal{G}_q$ is n-cell implementable, if and only if*

$$q \le n + 1.$$

Proof. Necessity will follow from the proof of Proposition 11 below and sufficiency follows by defining $x \mapsto \boldsymbol{u}(x)$ and $t \mapsto \boldsymbol{\vartheta}(t)$ as follows:

$$u_j(x) = \begin{cases} 0 & \text{if } x \le j \\ 1 & \text{if } x > j \end{cases} \quad \text{and} \quad \vartheta_j(t) = \begin{cases} * & \text{if } t \le j \\ 1 & \text{if } t > j \end{cases}$$

(these mappings are also shown in Table IV). $\square$

TABLE IV
Mappings for $\mathcal{G}_q$ that attain $q = n + 1$ under Scenario ($\circ *$).

x	$\boldsymbol{u}(x)$	t	$\boldsymbol{\vartheta}(t)$
0	$00000\ldots0$	0	$*****\ldots*$
1	$10000\ldots0$	1	$1****\ldots*$
2	$11000\ldots0$	2	$11***\ldots*$
3	$11100\ldots0$	3	$111**\ldots*$
$\vdots$	$\ddots$	$\vdots$	$\ddots$
$n{-}1$	$1111\ldots10$	$n{-}1$	$1111\ldots1*$
n	$11111\ldots1$	n	$11111\ldots1$

Proposition 11. *Under Scenario* ($**$)*, the family $\mathcal{G}_q$ is n-cell implementable, if and only if*

$$q \le \begin{cases} 2n & \text{if } n = 1, 2 \\ 2n+1 & \text{if } n \ge 3 \end{cases}.$$

Proof. Starting with proving necessity, suppose that Eq. (2) holds. For each $j \in [n\rangle$, let y_j denote the smallest $t \in [q\rangle$ such that $\vartheta_j(t) \in \mathbb{B}$ (define $y_j = \infty$ if no such t exists). By possibly switching between the roles of 0 and 1 in $t \mapsto \vartheta_j(t)$ and in $x \mapsto u_j(x)$, we can assume without loss of generality that $\vartheta_j(y_j) = 1$. We now observe from Eq. (2) that $u_j(x) \in \{1, *\}$ for every $x \geq y_j$. Thus, (2) still holds if we re-define $\vartheta_j(t)$ to be equal to 1 at every $t > y_j$ for which $\vartheta_j(t) = *$. Hence, we assume hereafter without loss of generality that

$$\vartheta_j(y_j) = 1 \;\text{ and }\; \vartheta_j(t) \in \mathbb{B} \;\text{ for every } t > y_j$$

(provided that $y_j < \infty$).

Next, for each $j \in [n\rangle$, we let z_j be the smallest $t \in [y_j : q\rangle$ such that $\vartheta_j(t) = 0$ (with $z_j = \infty$ if no such t exists). Note that for Eq. (2) to hold, we must have

$$u_j(x) = * \;\text{ for every } x \geq z_j. \tag{7}$$

In particular, under Scenario $(\circ *)$ (as in Proposition 10), we must have $z_j = \infty$ for every $j \in [n\rangle$.

In summary, the mapping $t \mapsto \boldsymbol{\vartheta}(t)$ is assumed hereafter to take the following form:

$$\vartheta_j(t) = \begin{cases} * & \text{if } t < y_j \\ 1 & \text{if } y_j \leq t < z_j \\ 0 & \text{if } t = z_j \\ 0 \text{ or } 1 & \text{if } t > z_j \end{cases}. \tag{8}$$

Next, we claim that

$$[1 : q\rangle \subseteq \bigcup_{j \in [n\rangle} \{y_j, z_j\} \tag{9}$$

(where the right-hand side is regarded as a set, ignoring multiplicities). Indeed, suppose to the contrary that there exists some $y \in [1 : q\rangle$ such that $y \notin \{y_j, z_j\}$ for every $j \in [n\rangle$. By (7) and (8) we then have, for every $j \in [n\rangle$:

$$\begin{array}{ll} \vartheta_j(y) = * & \text{if } y < y_j \\ \vartheta_j(y) = \vartheta_j(y-1) \; (= 1) & \text{if } y_j < y < z_j \\ u_j(y-1) = * & \text{if } y > z_j \end{array}. \tag{10}$$

Hence,[6]

$$\begin{aligned} \Gamma_q(y-1, y) &\overset{(2)}{=} \mathrm{T}(\boldsymbol{u}(y-1), \boldsymbol{\vartheta}(y)) \\ &\overset{y \neq y_j, z_j}{=} \Big(\bigwedge_{j \,:\, y < y_j} \mathrm{T}(u_j(y-1), \underbrace{\vartheta_j(y)}_{*}) \Big) \\ &\qquad \wedge \Big(\bigwedge_{j \,:\, y_j < y < z_j} \mathrm{T}(u_j(y-1), \vartheta_j(y)) \Big) \\ &\qquad \wedge \Big(\bigwedge_{j \,:\, y > z_j} \mathrm{T}(\underbrace{u_j(y-1)}_{*}, \vartheta_j(y)) \Big) \\ &\overset{(10)}{=} \bigwedge_{j \,:\, y_j < y < z_j} \mathrm{T}(u_j(y-1), \vartheta_j(y)) \\ &\overset{(10)}{=} \bigwedge_{j \,:\, y_j < y < z_j} \mathrm{T}(u_j(y-1), \vartheta_j(y-1)) \\ &\overset{(2)}{=} 1, \end{aligned}$$

which is a contradiction. By (9) we thus conclude that $q-1 \leq 2n$, thereby proving the necessary condition for $n \geq 3$ (leaving the special cases of $n = 1, 2$ to Appendix A). Moreover, when the range of $\boldsymbol{u}$ is constrained to $\mathbb{B}^n$, then $z_j = \infty$ for every $j \in [n\rangle$ and, so, by (9) we get $q - 1 \leq n$, thus proving the necessary condition in Proposition 10.

Sufficiency follows from the mappings shown in Table V; in that table, $\boldsymbol{u}(0)$ and $\boldsymbol{\vartheta}(2n)$ can be set to any two words in $\mathbb{B}_*^n$ that start and end with a 0 and $\mathrm{T}(\boldsymbol{u}(x), \boldsymbol{\vartheta}(2n)) = 0$ (e.g., we can take $\boldsymbol{u}(0)$ and $\boldsymbol{\vartheta}(2n)$ to be distinct in $\mathbb{B}^n$ that start and end with a 0). Note that this is always possible when $n \geq 3$; for $n = 1, 2$ we restrict the table to the rows that correspond to $x, t \in [2n\rangle$. $\square$

TABLE V

MAPPINGS FOR $\mathcal{G}_q$ THAT ATTAIN $q = 2n + 1$ UNDER SCENARIO $(**)$.

x	$\boldsymbol{u}(x)$	t	$\boldsymbol{\vartheta}(t)$
0	$0000\ldots0$	0	$****\ldots*$
1	$1000\ldots0$	1	$1***\ldots*$
2	$1100\ldots0$	2	$11**\ldots*$
$\vdots$	$\ddots$	$\vdots$	$\ddots$
$n{-}1$	$111\ldots10$	$n{-}1$	$111\ldots1*$
n	$1111\ldots1$	n	$1111\ldots1$
$n{+}1$	$*111\ldots1$	$n{+}1$	$0111\ldots1$
$n{+}2$	$**11\ldots1$	$n{+}2$	$0011\ldots1$
$\vdots$	$\ddots$	$\vdots$	$\ddots$
$2n{-}1$	$***\ldots*1$	$2n{-}1$	$000\ldots01$
$2n$	$****\ldots*$	$2n$	$0100\ldots0$

Proposition 12. *Under each of the scenarios $(*\bullet)$ or $(\bullet\bullet)$, the family $\mathcal{G}_q$ is n-cell implementable, if and only if*

$$q \leq 2n + 1.$$

Proof. Necessity will follow from the proof of Proposition 16 below, and sufficiency follows from the mappings $x \mapsto \boldsymbol{u}(x)$ and $t \mapsto \boldsymbol{\vartheta}(t)$ defined in (4)–(5) and extended to the domain $[2n + 1\rangle$ by:

$$\boldsymbol{u}(2n) = **\ldots* \quad \text{and} \quad \boldsymbol{\vartheta}(2n) = \bullet\bullet\ldots\bullet.$$

(For $n \geq 3$, we can also use the mappings in Table V.) $\square$

V. The family $\mathcal{G}_q \cup \mathcal{L}_q$

In this section, we consider the family $\Phi = \mathcal{G}_q \cup \mathcal{L}_q$. For the subset $\mathcal{G}_q$ of functions in this family, we use—as in Section IV—the notation $t \mapsto \boldsymbol{\vartheta}(t)$ for $\Gamma_q(\cdot, t) \mapsto \boldsymbol{\vartheta}(\Gamma_q(\cdot, t))$. For the remaining subset $\mathcal{L}_q$, we use $t \mapsto \boldsymbol{\vartheta}'(t)$ for $\mathrm{L}_q(\cdot, t) \mapsto \boldsymbol{\vartheta}(\mathrm{L}_q(\cdot, t))$ so that (2) becomes

$$\mathrm{L}_q(x, t) = \mathrm{T}(\boldsymbol{u}(x), \boldsymbol{\vartheta}'(t)). \tag{11}$$

Proposition 13. *Under Scenario $(*\circ)$, the family $\mathcal{G}_q \cup \mathcal{L}_q$ is n-cell implementable, if and only if*

$$q \leq n.$$

Proof. Sufficiency follows from Proposition 1. As for necessity, we consider first just the q mappings $x \mapsto \Gamma_q(x, t)$ in $\mathcal{G}_q \cup \mathcal{L}_q$ and refer to the containment (9) in the proof of Proposition 11. For Scenario $(*\circ)$ we have $y_j = 0$ for all $j \in [n\rangle$ and, so,

$$[1 : q\rangle \subseteq \big\{ z_j \,:\, j \in [n\rangle \big\}.$$

[6]A conjunction over an empty set is defined to be 1.

If $z_j = \infty$ for at least one $j \in [n\rangle$ we are done. Otherwise, by (7), we must have $\boldsymbol{u}(q-1) = **\ldots*$, yet this would imply $\mathrm{L}_q(q-1,0) = \mathrm{T}(\boldsymbol{u}(q-1), \boldsymbol{\vartheta}'(0)) = 1$, which is impossible. □

Proposition 14. *Under Scenario* $(\circ*)$*, the family* $\mathcal{G}_q \cup \mathcal{L}_q$ *is* n*-cell implementable, if and only if*

$$q \le n+1.$$

Proof. Necessity follows from Proposition 10. Sufficiency follows by taking the mappings $x \mapsto \boldsymbol{u}(x)$ and $t \mapsto \boldsymbol{\vartheta}(t)$ as in the proof of that proposition and defining the mapping $t \mapsto \boldsymbol{\vartheta}'(t)$ by

$$\vartheta'_j(t) = \begin{cases} 0 & \text{if } t \le j \\ * & \text{if } t > j \end{cases}.$$

□

Proposition 15. *Under each of the scenarios* $(**)$ *or* $(\bullet*)$*, the family* $\mathcal{G}_q \cup \mathcal{L}_q$ *is* n*-cell implementable, if and only if*

$$q \le \begin{cases} 2 & \text{if } n = 1 \\ 2n-1 & \text{if } n \ge 2 \end{cases}.$$

Proof. Sufficiency follows from Table VI for $n \ge 2$ (some symbols in the table are underlined to make it easier to see the general pattern). As for $n = 1$, we take:

x	$u(x)$
0	0
1	1

t	$\vartheta(t)$	$\vartheta'(t)$
0	$*$	0
1	1	$*$

We show necessity by induction on n. Suppose that Eq. (2) holds and, for $j \in [n\rangle$, let y_j and z_j be defined for the mapping $t \mapsto \boldsymbol{\vartheta}(t)$ in as in the proof of Proposition 11. Also, let y'_j be the largest $t \in [q\rangle$ such that $\vartheta'_j(t) \in \mathbb{B}$ (define $y'_j = -\infty$ if no such t exists) and let z'_j be the largest $t \in [y'_j\rangle$ such that $\vartheta'_j(t) \ne \vartheta'_j(y'_j)$ (with $z'_j = -\infty$ if no such t exists).[7] Denote by Y (respectively, Y') the set of all indexes $j \in [n\rangle$ such that y_j (respectively, y'_j) is finite and by Z $(\subseteq Y)$ (respectively, Z' $(\subseteq Y')$) the set of all indexes $j \in [n\rangle$ such that z_j (respectively, z'_j) is finite Clearly, $|Z| \le |Y| \le n$ and $|Z'| \le |Y'| \le n$ and, by (9) (when applied to $x \mapsto \boldsymbol{u}(x)$ and $t \mapsto \boldsymbol{\vartheta}(t)$ on the one hand, and to $x \mapsto \boldsymbol{u}(q-1-x)$ and $t \mapsto \boldsymbol{\vartheta}'(q-1-t)$ on the other hand), we get:

$$q-1 \le \min\{|Y|+|Z|, |Y'|+|Z'|\}. \tag{12}$$

If $Y = Z = [n\rangle$ then necessarily $\boldsymbol{u}(q-1) = **\ldots*$ (by (7)), yet this would imply by (2) that $\mathrm{L}_q(q-1,0) = 1$, which is impossible. Hence, $|Y|+|Z| \le 2n-1$ and, similarly, $|Y'|+|Z'| \le 2n-1$. We conclude that the right-hand side of (12) is at most $2n-1$, thereby establishing the necessity condition for the induction base $n = 1$, and also when $\min\{|Y|+|Z|, |Y'|+|Z'|\} \le 2n-2$.

It remains to consider the case where $Y = Y' = [n\rangle$ and $|Z| = |Z'| = n-1$. By (7), the entries of $\boldsymbol{u}(q-1)$ are then all $*$ except one, and the same applies to $\boldsymbol{u}(0)$. Without loss of generality we can assume that $Z = [n-1\rangle$, thereby implying that

$$\boldsymbol{u}(q-1) = **\ldots*\bullet \quad \text{or} \quad \boldsymbol{u}(q-1) = **\ldots*1.$$

In the first case we must have $\vartheta_{n-1}(t) = *$ for all $t \in [q\rangle$, thus effectively reducing the value of n by 1 when writing (2) for the functions in $\mathcal{G}_q$ within $\mathcal{G}_q \cup \mathcal{L}_q$; by Proposition 11 we then get $q \le 2(n-1)+1 = 2n-1$. In the remaining part of the proof we assume that $\boldsymbol{u}(q-1) = **\ldots*1$; this, in turn, forces having

$$\vartheta'_{n-1}(t) = 0 \ \text{ for all } t \in [q-1\rangle. \tag{13}$$

We distinguish between the two possible values for $\vartheta'_{n-1}(q-1)$.

Case 1: $\vartheta'_{n-1}(q-1) = 1$. Here we get from (13) that $u_{n-1}(x) = *$ for all $x \in [q-1\rangle$, which means that Eq. (2) holds for all $f \in \mathcal{G}_q \cup \mathcal{L}_q$ with q and n in (2) replaced by $q-1$ and $n-1$, respectively. Hence, by the induction hypothesis,

$$q-1 \le \max\{2(n-1)-1, 2\},$$

namely, $q \le \max\{2n-2, 3\} \le 2n-1$ when $n \ge 2$.

Case 2: $\vartheta'_{n-1}(q-1) = *$. Here we get from (13) that $Z' = Z = [n-1\rangle$, which means that the unique non-$*$ entry in $\boldsymbol{u}(0)$ must be its last (i.e., co-located with the unique non-$*$ entry in $\boldsymbol{u}(q-1)$). Recalling that $\vartheta'_{n-1}(0) = 0$ (by (13)), we must then have $\boldsymbol{u}(0) = **\ldots*0$, thereby implying that

$$\vartheta_{n-1}(t) = 1 \ \text{ for all } t \in [1:q\rangle.$$

Combining this with (13) we must also have $u_{n-1}(x) = *$ for all $x \in [1:q-1\rangle$. We conclude that Eq. (2) holds for all $f \in \mathcal{G}_q \cup \mathcal{L}_q$ with n in that equation replaced by $n-1$ and with x and t restricted to $[1:q-1\rangle$. Replacing x and t in (2) by $x+1$ and $t+1$, respectively, and noting that $\Gamma_{q-2}(x,t) = \Gamma_q(x+1,t+1)$ and $\mathrm{L}_{q-2}(x,t) = \mathrm{L}_q(x+1,t+1)$ for $x, t \in [q-2\rangle$, we get from the induction hypothesis that

$$q-2 \le \max\{2(n-1)-1, 2\},$$

namely, $q \le \max\{2n-1, 4\}$. This completes the proof for Case 2 when $n > 2$. As for $n = 2$, if $q = 4$ were attainable, then, from what we have just shown, we would have $u_1(1) = u_1(2) = *$, which would imply that $u_0(1) = \vartheta'_0(1) \in \mathbb{B}$ and $u_0(2) = \vartheta_0(2) \in \mathbb{B}$ and $u_0(1) \ne u_0(2)$. However, this would mean that $\mathrm{L}_4(x,0) = \mathrm{T}(u_0(x), \vartheta'_0(0)) = 1$ for at least one $x \in \{1,2\}$ (regardless of the value of $\vartheta'_0(0)$), which is a contradiction. □

Proposition 16. *Under each of the scenarios* $(*\bullet)$ *or* $(\bullet\bullet)$*, the family* $\mathcal{G}_q \cup \mathcal{L}_q$ *is* n*-cell implementable, if and only if*

$$q \le 2n.$$

Proof. Sufficiency follows from Proposition 3.

Necessity follows from arguments that are similar to (and in fact are even simpler than) those that we used in the proofs of Propositions 11 and 15. Specifically, for each $j \in [n\rangle$, let y_j denote the smallest $t \in [q\rangle$ such that $\vartheta_j(t) \in \mathbb{B} \cup \{\bullet\}$ (with $y_j = \infty$ if no such t exists). By the very same reasoning that was given in the proof of Proposition 11, we can assume

[7] In other words, if $\hat{y}_j$ and $\hat{z}_j$ are defined for the mapping $t \mapsto \hat{\vartheta}_j(t) = \vartheta'_j(q-1-t)$ as y_j and z_j were defined for $t \mapsto \vartheta_j(t)$ in the proof of Proposition 11, then $y'_j = q-1-\hat{y}_j$ and $z'_j = q-1-\hat{z}_j$.

TABLE VI
MAPPINGS FOR $\mathcal{G}_q \cup \mathcal{L}_q$ THAT ATTAIN $q = 2n-1$ UNDER SCENARIO $(**)$.

x	$\boldsymbol{u}(x)$
0	$\underline{00}****\ldots*$
1	$\underline{10}****\ldots*$
2	$\underline{1}*0***\ldots*$
3	$\underline{1}**0**\ldots*$
$\vdots$	$\vdots \quad \ddots$
$n-2$	$\underline{1}***\ldots*0*$
$n-1$	$\underline{1}****\ldots*0$
n	$\underline{11}****\ldots*$
$n+1$	$\underline{*1}1***\ldots*$
$n+2$	$\underline{*1}*1**\ldots*$
$n+3$	$\underline{*1}**1*\ldots*$
$\vdots$	$\vdots \quad \ddots$
$2n-3$	$\underline{*1}**\ldots*1*$
$2n-2$	$\underline{*1}***\ldots*1$

t	$\boldsymbol{\vartheta}(t)$	$\boldsymbol{\vartheta}'(t)$
0	$\underline{*}*****\ldots*$	$\underline{00}0000\ldots0$
1	$\underline{1}*****\ldots*$	$\underline{*0}1111\ldots1$
2	$\underline{1}1****\ldots*$	$\underline{*0}0111\ldots1$
3	$\underline{1}11***\ldots*$	$\underline{*0}0011\ldots1$
$\vdots$	$\ddots$	$\ddots$
$n-2$	$\underline{1}111\ldots1**$	$\underline{*0}000\ldots01$
$n-1$	$\underline{1}1111\ldots1*$	$\underline{*0}0000\ldots0$
n	$\underline{1}11111\ldots1$	$\underline{*}*0000\ldots0$
$n+1$	$\underline{01}1111\ldots1$	$\underline{*}**000\ldots0$
$n+2$	$\underline{01}0111\ldots1$	$\underline{*}***00\ldots0$
$n+3$	$\underline{01}0011\ldots1$	$\underline{*}****0\ldots0$
$\vdots$	$\ddots$	$\ddots$
$2n-3$	$\underline{01}00\ldots011$	$\underline{*}****\ldots*0$
$2n-2$	$\underline{01}000\ldots01$	$\underline{*}*****\ldots*$

hereafter without loss of generality that $\vartheta_j(y_j) \in \{1, \bullet\}$ and that $\vartheta_j(t) \in \mathbb{B} \cup \{\bullet\}$ for every $t > y_j$.

Next, for each $j \in [n\rangle$, let z_j be the smallest $t \in [y_j : q\rangle$ such that $\vartheta_j(t) \in \{0, \bullet\}$ (with $z_j = \infty$ if no such t exists). By (7) we then must have $u_j(x) = *$ for every $x \geq z_j$ and, so, without loss of generality we can assume that $\vartheta_j(t) = \bullet$ for every $t \geq z_j$. In summary, the mapping $t \mapsto \boldsymbol{\vartheta}(t)$ takes the following form:

$$\vartheta_j(t) = \begin{cases} * & \text{if } t < y_j \\ 1 & \text{if } y_j \leq t < z_j \\ \bullet & \text{if } t \geq z_j \end{cases} . \tag{14}$$

As our next step, we show that (9) holds. Indeed, if there were $y \in [1 : q\rangle$ that did not belong to the right-hand side of (9) then, from (14), we would have $\boldsymbol{\vartheta}(y) = \boldsymbol{\vartheta}(y-1)$. Yet, by (2), this would mean that $\Gamma_q(y-1, y) = \Gamma_q(y, y) = 1$, which is a contradiction. From (9) we now get the necessary condition in Proposition 12 (for the family $\Phi = \mathcal{G}_q$). The necessary condition for $\Phi = \mathcal{G}_q \cup \mathcal{L}_q$ follows by showing that $z_j = \infty$ for at least one $j \in [n\rangle$ (and, therefore, the right-hand side of (9) contains less than $2n$ elements). Otherwise, we would have $\boldsymbol{\vartheta}(q-1) = \bullet\bullet\ldots\bullet$ and, consequently, $\boldsymbol{u}(q-1) = **\ldots*$, which, with (11), would yield $\mathrm{L}_q(q-1, 0) = 1$, thereby reaching a contradiction. $\square$

VI. THE WHOLE SET $\mathcal{F}_q$

In this section, we treat the case where $\Phi = \mathcal{F}_q$ (the whole set of functions $[q\rangle \to \mathbb{B}$). Differently from previous sections, we start with Scenarios $(\bullet*)$ and $(\bullet\bullet)$, as they are rather straightforward.

Proposition 17. *Under each of the scenarios* $(*\bullet)$ *or* $(\bullet\bullet)$*, the set* $\mathcal{F}_q$ *is* n*-cell implementable, if and only if*

$$q \leq 2n.$$

Proof. Sufficiency follows from Proposition 3 and necessity follows either from Proposition 9 or by a counting argument: $\boldsymbol{\vartheta} : \mathcal{F}_q \to \mathbb{B}_\bullet^n$ is injective and, so, $2^q \leq 2^{2n}$. $\square$

The next proposition covers all the remaining scenarios in Table II.

Proposition 18. *Under each of the scenarios* $(*\circ)$*,* $(\circ*)$*,* $(**)$*, or* $(\bullet*)$*, the set* $\mathcal{F}_q$ *is* n*-cell implementable, if and only if*

$$q \leq n.$$

Sufficiency follows from Proposition 1. As for necessity, for Scenarios $(*\circ)$ and $(\circ*)$ it is implied by Proposition 8 (for Scenario $(*\circ)$ we can also use Proposition 13 or just a counting argument). For Scenarios $(**)$ and $(\bullet*)$, however, some more effort is needed; notice that a counting argument only leads to $2^q \leq 3^n$, namely, to the weaker inequality $q \leq n \cdot \log_2 3$.

The proof of Proposition 18 will use the following notation and lemma. Recalling the partial ordering $\preceq$ of Section II, for elements $u, v \in \mathbb{B}_\bullet$, we denote by $\mu(u, v)$ the largest element $s \in \mathbb{B}_\bullet$ such that both $s \preceq u$ and $s \preceq v$, where "largest" is with respect to $\preceq$. Thus, for every $u \in \mathbb{B}_\bullet$,

$$\mu(u, \bullet) = \bullet, \quad \mu(u, *) = \mu(u, u) = u, \quad \text{and} \quad \mu(0, 1) = \bullet.$$

The next lemma is easily verified.

Lemma 19. *For every* $u, v, \vartheta \in \mathbb{B}_\bullet$*,*

$$\mathrm{T}(\mu(u, v), \vartheta) = \mathrm{T}(u, \vartheta) \wedge \mathrm{T}(v, \vartheta).$$

Proof of Proposition 18 (necessity). We prove necessity under (the loosest) Scenario $(\bullet*)$ by induction on q, with the induction base ($q = 2$) following from a simple counting argument: there are four distinct functions $f : [2\rangle \to \mathbb{B}$ yet only three choices for (the scalar) $\vartheta(f)$ in this case, hence we must have $n \geq 2$.[8]

Turning to the induction step, suppose to the contrary that $\mathcal{F}_q$ is n-cell implementable for $q > n$ and let $\boldsymbol{u} : [q\rangle \to \mathbb{B}_\bullet^n$ and $\boldsymbol{\vartheta} : \mathcal{F}_q \to \mathbb{B}_*^n$ be mappings such that (2) holds for all the functions in $\mathcal{F}_q$; obviously, both these mappings are injective.

Consider the images of the functions in $\mathcal{N}_q$ under $\boldsymbol{\vartheta}$. As argued in the proof of Proposition 8, we can assume that for each $t \in [q\rangle$ there is a unique entry in $\boldsymbol{\vartheta}(\mathrm{N}_q(\cdot, t))$ which is non-$*$. From $q > n$ we get, by the pigeonhole principle, that there exist two distinct elements $t_0, t_1 \in [q\rangle$ for which the position j of such an entry is the same, say $j = n-1$. Without loss of generality we can assume further that $t_0 = q-2$, $t_1 = q-1$,

$$\vartheta_{n-1}(\mathrm{N}_q(\cdot, q-2)) = 0, \quad \text{and} \quad \vartheta_{n-1}(\mathrm{N}_q(\cdot, q-1)) = 1.$$

This, in turn, implies that for every $x \in [q\rangle$:

$$u_{n-1}(x) = \begin{cases} * & \text{if } x \in [q-2\rangle \\ 1 & \text{if } x = q-2 \\ 0 & \text{if } x = q-1 \end{cases} . \tag{15}$$

Next, we define the mappings

$$\boldsymbol{u}' : [q-1\rangle \to \mathbb{B}_\bullet^{n-1} \quad \text{and} \quad \boldsymbol{\vartheta}' : \mathcal{F}_{q-1} \to \mathbb{B}_*^{n-1}$$

[8] Note that $n = 1$ would suffice for $q = 2$ if we allowed $\vartheta(f)$ to take the value $\bullet$ as well: this corresponds to Scenario $(\circ\bullet)$ and it differs from Scenario $(\circ*)$ only when $n = 1$ and $q = 2$.

as follows: for every $x \in [q-1\rangle$ and $j \in [n-1\rangle$,

$$u'_j(x) = \begin{cases} u_j(x) & \text{if } x \in [q-2\rangle \\ \mu(u_j(q-2), u_j(q-1)) & \text{if } x = q-2 \end{cases}, \quad (16)$$

and for every $f \in \mathcal{F}_{q-1}$ and $j \in [n-1\rangle$,

$$\vartheta'_j(f) = \vartheta_j(\tilde{f}), \quad (17)$$

where $\tilde{f}$ is the extension of f to the domain $[q\rangle$ with

$$\tilde{f}(q-1) = f(q-2).$$

In particular, it follows from (16), (17), and Lemma 19 that for every $j \in [n-1\rangle$:

$$\begin{aligned} &\mathrm{T}(u'_j(q-2), \vartheta'_j(f)) \\ &\quad = \mathrm{T}(u_j(q-2), \vartheta_j(\tilde{f})) \wedge \mathrm{T}(u_j(q-1)), \vartheta_j(\tilde{f})). \end{aligned} \quad (18)$$

We show that

$$f(x) = \mathrm{T}(\boldsymbol{u}'(x), \boldsymbol{\vartheta}'(f))$$

for every $x \in [q-1\rangle$ and $f \in \mathcal{F}_{q-1}$ which, in turn, will imply that $\mathcal{F}_{q-1}$ is $(n-1)$-cell implementable, thereby contradicting the induction hypothesis. We distinguish between three cases.

Case 1: $x \in [q-2\rangle$. We recall from (15) that $u_{n-1}(x) = *$ and, so,

$$f(x) = \tilde{f}(x) \overset{(2)}{=} \mathrm{T}(\boldsymbol{u}(x), \boldsymbol{\vartheta}(\tilde{f})) \overset{(16)+(17)}{=} \mathrm{T}(\boldsymbol{u}'(x), \boldsymbol{\vartheta}'(f)).$$

Case 2: $x = q-2$ *and* $f(q-2) = 1$. Here $\tilde{f}(q-2) = \tilde{f}(q-1) = 1$ and, therefore, from (2), for every $j \in [n\rangle$,

$$\mathrm{T}(u_j(q-2), \vartheta_j(\tilde{f})) = \mathrm{T}(u_j(q-1), \vartheta_j(\tilde{f})) = 1.$$

Hence, by (18), for every $j \in [n-1\rangle$,

$$\mathrm{T}(u'_j(q-2), \vartheta'_j(f)) = 1,$$

namely, $\mathrm{T}(\boldsymbol{u}'(q-2), \boldsymbol{\vartheta}'(f)) = 1 = f(q-2)$.

Case 3: $x = q-2$ *and* $f(q-2) = 0$. Here $\tilde{f}(q-2) = \tilde{f}(q-1) = 0$; yet, from (15) we have $\mathrm{T}(u_{n-1}(y), \vartheta_{n-1}(\tilde{f})) = 1$ for at least one $y \in \{q-2, q-1\}$. Hence, from $\mathrm{T}(\boldsymbol{u}(y), \boldsymbol{\vartheta}(\tilde{f})) = \tilde{f}(y) = 0$ there must be at least one index $j \in [n-1\rangle$ for which

$$\mathrm{T}(u_j(y), \vartheta_j(\tilde{f})) = 0,$$

and for that index we have, by (18),

$$\mathrm{T}(u'_j(q-2), \vartheta'_j(f)) = 0,$$

namely, $\mathrm{T}(\boldsymbol{u}'(q-2), \boldsymbol{\vartheta}'(f)) = 0 = f(q-2)$. $\square$

A. Connection to the VC dimension of Boolean monomials

Proposition 18 can be stated also in terms of the VC dimension of the following collection, $\mathcal{H}_n$, of 3^n subsets of $\mathbb{B}_\bullet^n$:

$$\mathcal{H}_n = \{\mathcal{S} = \mathcal{S}(\boldsymbol{\vartheta}) : \boldsymbol{\vartheta} \in \mathbb{B}_*^n\},$$

where, for each $\boldsymbol{\vartheta} \in \mathbb{B}_*^n$,

$$\mathcal{S}(\boldsymbol{\vartheta}) = \{\boldsymbol{v} \in \mathbb{B}_\bullet^n : \mathrm{T}(\boldsymbol{v}, \boldsymbol{\vartheta}) = 1\}.$$

We demonstrate this next.

Let $\boldsymbol{u} : [q\rangle \to \mathbb{B}_\bullet^n$ and $\boldsymbol{\vartheta} : \mathcal{F}_q \to \mathbb{B}_*^n$ be injective mappings. The following two conditions are equivalent for any $f \in \mathcal{F}_q$.

- Eq. (2) holds for f.
- The images of $x \mapsto \boldsymbol{u}(x)$ form a subset $\mathcal{U} \subseteq \mathbb{B}_\bullet^n$ of size q such that

$$\mathcal{U} \cap \mathcal{S}(\boldsymbol{\vartheta}(f)) = \{\boldsymbol{u}(x) : x \in [q\rangle \text{ such that } f(x) = 1\}. \quad (19)$$

In particular, if Eq. (2) holds for every $f \in \mathcal{F}_q$, then (19) implies that $\mathcal{U}$ is *shattered* by $\mathcal{H}_n$: each of the 2^q subsets of $\mathcal{U}$ can be expressed as an intersection $\mathcal{U} \cap \mathcal{S}$, for some $\mathcal{S} \in \mathcal{H}_n$. Conversely, if $\mathcal{U}$ is a subset of size q of $\mathbb{B}_\bullet^n$ that is shattered by $\mathcal{H}_n$, we can fix some arbitrary bijection $\boldsymbol{u} : [q\rangle \to \mathcal{U}$ and define a mapping $\boldsymbol{\vartheta} : \mathcal{F}_q \to \mathbb{B}_*^n$ so that (19) holds for every $f \in \mathcal{F}_q$; specifically, we select $\boldsymbol{\vartheta}(f)$ to be such that $\mathcal{S}(\boldsymbol{\vartheta}(f))$ is an element $\mathcal{S} \in \mathcal{H}_n$ for which $\mathcal{U} \cap \mathcal{S}$ equals the subset of $\mathcal{U}$ given by the right-hand side of (19). The largest size q of $\mathcal{U}$ for which this holds is the VC dimension of $\mathcal{H}_n$ [2, §7.3]. Thus, Proposition 18 states that the VC dimension of $\mathcal{H}_n$ is n.

Equivalently, we can state Proposition 18 in terms of the VC dimension of the collection of n-variate Boolean monomials where the evaluation points are taken from $\mathbb{B}_\bullet^n$ (rather than just from $\mathbb{B}^n$). Specifically, given a vector of n Boolean indeterminates, $\boldsymbol{\xi} = (\xi_j)_{j \in [n\rangle}$, we associate with every word $\boldsymbol{\vartheta} = (\vartheta_j)_{j \in [n\rangle} \in \mathbb{B}_*^n$ the n-variate Boolean monomial

$$M_{\boldsymbol{\vartheta}}(\boldsymbol{\xi}) = \Big(\bigwedge_{j \,:\, \vartheta_j = 0} \xi_j \Big) \wedge \Big(\bigwedge_{j \,:\, \vartheta_j = 1} \overline{\xi}_j \Big),$$

where $\overline{\xi}_j$ stands for the complement of ξ_j. Substituting an element of $\mathbb{B}$ into a variable ξ_j carries its ordinary meaning (with $\overline{0} = 1$ and $\overline{1} = 0$), whereas when substituting $*$ (respectively, $\bullet$), both ξ_j and $\overline{\xi}_j$ are defined to be 1 (respectively, 0). Under these rules, $\mathcal{S}(\boldsymbol{\vartheta})$ is the set of all words in $\mathbb{B}_\bullet^n$ at which $M_{\boldsymbol{\vartheta}}(\boldsymbol{\xi})$ evaluates to 1. Scenario ($\circ *$) corresponds to the case where the evaluation points are restricted to $\mathbb{B}^n$ and, for this scenario, it was shown in [11] that the VC dimension of the set of all n-variate monomials equals n. Proposition 18 implies that the VC dimension does not increase even if we extend the set of evaluation points to $\mathbb{B}_\bullet^n$.[9]

B. Subsets that are as hard to implement as $\mathcal{F}_q$

Recall that under Scenarios ($*\circ$), ($\circ *$), ($*\bullet$), and ($\bullet\bullet$), there are small subsets of $\mathcal{F}_q$ which are n-cell implementable, (if and) only if $\mathcal{F}_q$ is; e.g., $\mathcal{N}_q$ is such as subset of size $q = \log_2 |\mathcal{F}_q|$. In contrast, it turns out that under Scenarios ($**$) and ($\bullet *$), the condition $q \le n$ becomes necessary only for fairly large subsets of $\mathcal{F}_q$. The next two propositions (which we prove in Appendix B) imply that, with very few exceptions, a deletion of just a single function from $\mathcal{F}_q$ results in a subset which is $(q-1)$-cell implementable.

Henceforth, $x \mapsto \mathit{1}_q(x)$ stands for the tautology function over $[q\rangle$ (which evaluates to 1 on all the elements of $[q\rangle$).

Proposition 20. *Let $\Phi = \mathcal{F}_q \backslash \{g\}$ where g is any function in $\mathcal{F}_q$ that is not in $\mathcal{E}_q \cup \{\mathit{1}_q\}$. Under Scenario ($**$), the subset Φ is n-cell implementable, if and only if*

$$q \le n+1.$$

[9] A setting where some variables are allowed to be "unspecified" has been studied in the literature (see, for example, [3]), yet the meaning of $*$ therein is different.

Proposition 21. *Let $\Phi = \mathcal{F}_q \setminus \{g\}$ where g is any function in $\mathcal{F}_q \setminus \{\mathit{1}_q\}$. Under Scenario ($\bullet *$), the subset Φ is n-cell implementable, if and only if*

$$q \leq n + 1.$$

It readily follows from Proposition 20 that under Scenario ($**$), any subset of $\mathcal{F}_q$ of size smaller than $2^q - q$ is $(q-1)$-cell implementable, except (possibly) for the subset

$$\mathcal{F}_q^\star = \mathcal{F}_q \setminus \left(\mathcal{E}_q \cup \{\mathit{1}_q\}\right). \tag{20}$$

And by the following result (which we also prove in Appendix B), this subset is indeed an exception.

Proposition 22. *For $q \geq 3$, the subset $\mathcal{F}_q^\star$ in (20) is n-cell implementable under Scenario ($**$), (if and) only if $\mathcal{F}_q$ is.*

Similar claims can be stated for Scenario ($\bullet *$), yet with the subset in (20) replaced by

$$\mathcal{F}_q \setminus \{\mathit{1}_q\}.$$

Specifically, all other proper subsets of $\mathcal{F}_q$ are $(q-1)$-cell implementable.

VII. Discussion

We can see from Table II that under all scenarios and for all families of functions therein except $\mathcal{E}_q$, the number n of cells must grow linearly with the alphabet size q. While this is expected for the whole set $\mathcal{F}_q$ (simply because of a counting argument), this also holds for the families $\mathcal{N}_q$, $\mathcal{G}_q$, $\mathcal{L}_q$ (and $\mathcal{G}_q \cup \mathcal{L}_q$), which are only of size $O(q)$. This means that for the latter families, CAM implementations are necessarily exponential in the (bit) representation of the parameter $t \in [q\rangle$ that identifies each function in the family, as well as in the representation of the input $x \in [q\rangle$ to each function. Thus, some hardware modification is inevitable if more efficient implementations are sought.

And this is indeed possible. For example, suppose that $q = 2n$ and we wish to implement each function $x \mapsto \mathrm{N}_{q^2}(x, t)$ in $\mathcal{N}_{q^2}$ under Scenario ($**$), ($\bullet *$), ($* \bullet$), or ($\bullet\bullet$). We can express each $x, t \in [q^2\rangle$ as

$$x = x_1 q + x_0 \quad \text{and} \quad t = t_1 q + t_0,$$

where $x_0, x_1, t_0, t_1 \in [q\rangle$, and observe that

$$\mathrm{N}_{q^2}(x, t) = \mathrm{N}_q(x_0, t_0) \vee \mathrm{N}_q(x_1, t_1). \tag{21}$$

Thus, while we have squared the alphabet size, the right-hand side of (21) requires only $q = 2n$ cells (and not $q^2/2 = 2n^2$ cells), yet such cells have to be arranged in two n-cell blocks that are connected *serially* (to compute the disjunction $\vee$), as opposed to the ordinary *parallel* connection among the cells along a row in the CAM.

Appendix A
Special cases in Proposition 11

We verify here the necessary condition $q \leq 2$ for $n = 1$, and $q \leq 4$ for $n = 2$.

Starting with $n = 1$, suppose to the contrary that Eq. (2) can hold for $q = 2n + 1 = 3$. By (9) we then have $y_0 = 1$ and $z_0 = 2$, namely, $\vartheta(0) = *$, $\vartheta(1) = 1$, and $\vartheta(2) = 0$. Yet this means that $\Gamma_3(0, t) = \mathrm{T}(u(0), \vartheta(t)) = 1$ for at least one $t \in \{1, 2\}$ (regardless of the value of $u(0)$), which is absurd.

Turning to $n = 2$, suppose to the contrary that Eq. (2) holds for $q = 2n + 1 = 5$. Again, by (9) we have $\{1, 2, 3, 4\} = \{y_0, z_0, y_1, z_1\}$, where $y_0 < z_0$ and $y_1 < z_1$. Without loss of generality we assume that $z_0 > z_1$, in which case $z_0 = 4$ and, so $\vartheta_0(3) = 1$ and $\vartheta_0(4) = 0$ (see (8)). Now, if $z_1 < 3$ then, by (7), we would have $u_1(2) = *$; yet then $\Gamma_5(2, t) = \mathrm{T}(u_0(2), \vartheta_0(t)) = 1$ for at least one $t \in \{3, 4\}$, which is impossible. Hence, $z_1 = 3$, and from (8) we get $\vartheta_1(2) = 1$ and $\vartheta_1(3) = 0$; moreover, $\{y_0, y_1\} = \{1, 2\}$ and, so, $\vartheta_0(2) = 1$. In summary, we have shown that $\boldsymbol{\vartheta}(2) = 11$, $\boldsymbol{\vartheta}(3) = 10$, and $\vartheta_0(4) = 0$. Now, since $\Gamma_5(x, t) = \mathrm{T}(\boldsymbol{u}(x), \boldsymbol{\vartheta}(t)) = 0$ for $x \in \{0, 1\}$ and $t \in \{2, 3\}$, we must have $u_0(0) = u_0(1) = 0$ $(= \vartheta_0(4))$, which implies that $u_1(0) \neq u_1(1)$. But this means that $\Gamma_5(x, 4) = \mathrm{T}(u_1(x), \vartheta_1(4)) = 1$ for at least one $x \in \{0, 1\}$, which is a contradiction.

Appendix B
More on the whole set $\mathcal{F}_q$

We provide in this appendix the proofs of Propositions 20, 21, and 22.

We will sometimes find it convenient to represent a function $f \in \mathcal{F}_q$ by its truth table $\boldsymbol{f} = (f(x))_{x \in [q\rangle} \in \mathbb{B}^q$; for example, $11\ldots1$ represents the tautology function $x \mapsto \mathit{1}_q(x)$ over $[q\rangle$ and $00\ldots0$ represents the all-zero function, which we denote by $x \mapsto \mathit{0}_q(x)$.

The next lemma provides the necessity part in Proposition 21 (and, therefore, also in Proposition 20).

Lemma 23. *Let $\Phi = \mathcal{F}_q \setminus \{g\}$ where g is any function in $\mathcal{F}_q \setminus \{\mathit{1}_q\}$. Under Scenario ($\bullet *$), the subset Φ is n-cell implementable, only if*

$$q \leq n + 1.$$

Proof. Suppose that (2) holds for the functions in Φ with the mappings $\boldsymbol{u}' : [q\rangle \to \mathbb{B}_\bullet^n$ and $\boldsymbol{\vartheta}' : \Phi \to \mathbb{B}_*^n$. Since $g \neq \mathit{1}_q$, there exists a function $\overline{g} \in \Phi$ which is identical to g on the whole domain $[q\rangle$ except for one element $y \in [q\rangle$ at which $g(y) = 0$ yet $\overline{g}(y) = 1$.

Define the mappings $\boldsymbol{u} : [q\rangle \to \mathbb{B}_\bullet^{n+1}$ and $\boldsymbol{\vartheta} : \mathcal{F}_q \to \mathbb{B}_*^{n+1}$ as follows. For every $x \in [q\rangle$,

$$\boldsymbol{u}(x) = \begin{cases} \boldsymbol{u}'(x)\, * & \text{if } x \neq y \\ \boldsymbol{u}'(y)\, 0 & \text{if } x = y \end{cases}$$

(namely, $\boldsymbol{u}(x)$ is obtained from $\boldsymbol{u}'(x)$ by appending a $*$ or a 0, depending on x), and for every $f \in \mathcal{F}_q$,

$$\boldsymbol{\vartheta}(f) = \begin{cases} \boldsymbol{\vartheta}'(f)\, * & \text{if } f \neq g \\ \boldsymbol{\vartheta}'(\overline{g})\, 1 & \text{if } f = g \end{cases}.$$

It can be verified that $f(x) = \mathrm{T}(\boldsymbol{u}(x), \boldsymbol{\vartheta}(f))$ for every $x \in [q\rangle$ and $f \in \mathcal{F}_q$, which means that $\mathcal{F}_q$ is $(n+1)$-cell implementable under Scenario ($\bullet *$). The result follows from Proposition 18. □

The next lemma provides the sufficiency part in Proposition 20 (and, therefore, also in Proposition 21 for $q \geq 3$) for the special case where $\boldsymbol{f}$ contains exactly one 0; due to symmetry, we can assume that $\boldsymbol{f} = 11\ldots110$, i.e., $f(x) = \mathrm{L}_q(x, q-2)$.

Lemma 24. *Under Scenario* $(**)$*, the subset* $\Phi = \mathcal{F}_q \setminus \{\mathrm{L}_q(\cdot, q-2)\}$ *is* n*-cell implementable, whenever*

$$3 \le q \le n+1.$$

Proof. We present mappings $\boldsymbol{u} : [q\rangle \to \mathbb{B}_*^{q-1}$ and $\boldsymbol{\vartheta} : \Phi \to \mathbb{B}_*^{q-1}$ for which (2) holds for every $f \in \Phi$. We take the mapping $\boldsymbol{u}$ to be

$$\boldsymbol{u}(q-1) = 00\ldots0 \tag{22}$$

and, for $x \in [q-1\rangle$:

$$u_j(x) = \begin{cases} 1 & \text{if } x = j \\ 0 & \text{if } x = \langle j{+}1\rangle \\ * & \text{otherwise} \end{cases}, \tag{23}$$

where $\langle\cdot\rangle = \langle\cdot\rangle_{q-1}$ denotes the remainder (in $[q-1\rangle$) modulo $q-1$; namely, $\boldsymbol{u}(x)$ is obtained from the unary representation of x by changing all the 0's into $*$'s, except the 0 that immediately precedes the (unique) 1 in that representation (with indexes extended cyclically modulo $q-1$). Table VII shows this mapping for $q = n+1 = 5$. Thus, $\boldsymbol{u}(x)$ contains at least one 0 for every $x \in [q\rangle$ (recall that $n \ge 2$).

Turning to the mapping $\boldsymbol{\vartheta}$, we select it so that

$$\boldsymbol{\vartheta}(\mathit{0}_q) = 11\ldots1 \tag{24}$$

(indeed, $\mathrm{T}(\boldsymbol{u}(x), \boldsymbol{\vartheta}(\mathit{0}_q)) = 0$ for all $x \in [q\rangle$); for all functions other than $\mathit{0}_q$ and $\mathrm{L}_q(\cdot, q-2)$ we let

$$\vartheta_j(f) = \begin{cases} 0 & \text{if } f(j) = 0 \\ 1 & \text{if } f(j) = 1 \text{ and } f(\langle j{+}1\rangle) = f(q{-}1) = 0 \\ * & \text{otherwise} \end{cases}. \tag{25}$$

Specifically, $\boldsymbol{\vartheta}(f)$ is obtained from the $(q-1)$-prefix of the truth table, $\boldsymbol{f}$, of f by changing into $*$'s all the 1's (if $f(q-1) = 1$), or just the 1's that immediately precede other 1's (if $f(q-1) = 0$, with indexes taken modulo $q-1$). Table VIII shows this mapping for $q = n+1 = 4$ (note the missing entry from the table, $\boldsymbol{f} = 1110$, which corresponds to $f(x) = \mathrm{L}_4(x, 2)$). Thus, $\boldsymbol{\vartheta}(f)$ will contain at least one 1 if and only if $f(q-1) = 0$ and, so, $\mathrm{T}(\boldsymbol{u}(q-1), \boldsymbol{\vartheta}(f)) = f(q-1)$.

It remains to show that Eq. (2) holds when $x \in [q-1\rangle$ and $f \ne \mathit{0}_q$. If x is such that $f(x) = 0$ then, by (23) and (25),

$$u_x(x) = 1 \quad \text{yet} \quad \vartheta_x(f) = 0$$

and, so, $\mathrm{T}(\boldsymbol{u}(x), \boldsymbol{\vartheta}(f)) = 0 = f(x)$. On the other hand, if $f(x) = 1$ then, for every $j \in [q-1\rangle$:

$$u_j(x) = * \quad \text{if } j \notin \{x, \langle x-1\rangle\}$$

and

$$u_j(x) = 1 \ \text{ and } \ \vartheta_j(f) \in \{1, *\} \quad \text{if } j = x.$$

As for the index $j = \langle x-1\rangle$, here $u_j(x) = 0$ while $\vartheta_j(f) \ne 1$ in this case (since $f(\langle j+1\rangle) = f(x) = 1$). Hence, $\mathrm{T}(\boldsymbol{u}(x), \boldsymbol{\vartheta}(f)) = 1 = f(x)$. □

TABLE VII
MAPPING $x \mapsto \boldsymbol{u}(x)$ FOR $q = n+1 = 5$ IN THE PROOF OF LEMMA 24.

x	$\boldsymbol{u}(x)$
0	1∗∗0
1	01∗∗
2	∗01∗
3	∗∗01
4	0000

TABLE VIII
MAPPING $f \mapsto \boldsymbol{\vartheta}(f)$ FOR $q = n+1 = 4$ IN THE PROOF OF LEMMA 24.

$\boldsymbol{f}$	$\boldsymbol{\vartheta}(f)$	$\boldsymbol{f}$	$\boldsymbol{\vartheta}(f)$
0000	111	0001	000
1000	100	1001	∗00
0100	010	0101	0∗0
1100	∗10	1101	∗∗0
0010	001	0011	00∗
1010	10∗	1011	∗0∗
0110	0∗1	0111	0∗∗
		1111	∗∗∗

Remark 2. The case $q = n+1 = 2$ was excluded from Lemma 24, since, under Scenario $(**)$, the lemma does not hold for the set

$$\Phi = \mathcal{F}_2 \setminus \{\mathrm{L}_2(\cdot, 0)\} = \{\mathit{0}_2, \Gamma_2(\cdot, 1), \mathit{1}_2\}. \tag{26}$$

Indeed, regardless of how we select $\vartheta(\mathit{0}_2)$, we would get $\mathrm{T}(u(x), \vartheta(\mathit{0}_2)) = 1$ for at least one $x \in [2\rangle$.

However, the lemma does hold for the set (26) under Scenario $(\bullet *)$ by taking

$$u(0) = \bullet, \quad u(1) = 1$$

and

$$\vartheta(\mathit{0}_2) = 0, \quad \vartheta(\Gamma_2(\cdot, 1)) = 1, \quad \vartheta(\mathit{1}_2) = *.$$

□

Proof of Proposition 20 (sufficiency). We show that $q = n{+}1$ is achievable by induction on n. For $n = 1$, the only choice for g is $\mathit{0}_2$, in which case we take

$$u(0) = 0, \quad u(1) = 1, \tag{27}$$

and

$$\vartheta(\mathrm{L}_2(\cdot, 0)) = 0, \quad \vartheta(\Gamma_2(\cdot, 1)) = 1, \quad \vartheta(\mathit{1}_2) = *. \tag{28}$$

Turning to the induction step, assume that $q = n+1 \ge 3$ and let $\boldsymbol{g} = (g(x))_{x \in [q\rangle}$ be the truth table of some function $g \in \mathcal{F}_q$ not in $\mathcal{E}_q \cup \{\mathit{1}_q\}$. Without loss of generality we can assume that

$$\boldsymbol{g} = \underbrace{11\ldots1}_{t+1}00\ldots0,$$

namely, $g(x) = \mathrm{L}_q(x, t)$, where $t \ne 0, q{-}1$. The case $t = q{-}2$ is covered by Lemma 24, so we assume hereafter that $t \le q{-}3$ (i.e., $g(q-2) = g(q-1) = 0$).

Consider the function $g' \in \mathcal{F}_{q-1}$ which is the restriction of g to the domain $[q-1\rangle$. Then $g' \notin \mathcal{E}_{q-1} \cup \{\mathit{1}_{q-1}\}$ by our assumptions and, so, by the induction hypothesis, there exist $\boldsymbol{u}' : [q-1\rangle \to \mathbb{B}_*^{n-1}$ and $\boldsymbol{\vartheta}' : \mathcal{F}_{q-1} \to \mathbb{B}_*^{n-1}$ such that

$$\mathrm{T}(\boldsymbol{u}'(x), \boldsymbol{\vartheta}'(f')) = f'(x) \tag{29}$$

for every $f' \in \mathcal{F}_{q-1} \setminus \{g'\}$ and $x \in [q-1\rangle$.

Define the mapping $\boldsymbol{u} : [q\rangle \to \mathbb{B}_*^n$ by

$$\boldsymbol{u}(q-1) = {*}{*}\ldots{*}{*}1 \tag{30}$$

and, for $x \in [q-1\rangle$:

$$\boldsymbol{u}(x) = \begin{cases} \boldsymbol{u}'(x)\,0 & \text{if } g(x) = 0 \\ \boldsymbol{u}'(x)\,{*} & \text{if } g(x) = 1 \end{cases} . \tag{31}$$

Let $\overline{g} : [q\rangle \to \mathbb{B}$ be the function which is identical to g except that $\overline{g}(q-1) = 1$. We define the mapping $\boldsymbol{\vartheta} : \mathcal{F}_q \setminus \{g\} \to \mathbb{B}_*^n$ by

$$\boldsymbol{\vartheta}(\overline{g}) = {*}{*}\ldots{*}{*}1 \tag{32}$$

while for $f \in \mathcal{F}_q \setminus \{g, \overline{g}\}$:

$$\boldsymbol{\vartheta}(f) = \begin{cases} \boldsymbol{\vartheta}'(f')\,0 & \text{if } f(q-1) = 0 \\ \boldsymbol{\vartheta}'(f')\,{*} & \text{if } f(q-1) = 1 \end{cases} , \tag{33}$$

where f' is the restriction of f to the domain $[q-1\rangle$.

We now verify that Eq. (2) holds for every $x \in [q\rangle$ and $f \in \mathcal{F}_q \setminus \{g\}$. For $x = q-1$ and $f = \overline{g}$ we have

$$\mathrm{T}(\boldsymbol{u}(q-1), \boldsymbol{\vartheta}(\overline{g})) \stackrel{(30)+(32)}{=} 1 = \overline{g}(q-1),$$

while for $x = q-1$ and $f \neq \overline{g}$,

$$\mathrm{T}(\boldsymbol{u}(q-1), \boldsymbol{\vartheta}(f)) \stackrel{(30)}{=} \mathrm{T}(1, \vartheta_{n-1}(f)) \stackrel{(33)}{=} f(q-1).$$

Assuming now that $x \in [q-1\rangle$, for $f = \overline{g}$,

$$\mathrm{T}(\boldsymbol{u}(x), \boldsymbol{\vartheta}(\overline{g})) \stackrel{(32)}{=} \mathrm{T}(u_{n-1}(x), 1) \stackrel{(31)}{=} g(x) = \overline{g}(x),$$

while for $f \neq \overline{g}$,

$$\mathrm{T}(\boldsymbol{u}(x), \boldsymbol{\vartheta}(f)) \stackrel{(31)+(33)}{=} \mathrm{T}(\boldsymbol{u}'(x), \boldsymbol{\vartheta}'(f')) \stackrel{(29)}{=} f'(x) = f(x).$$

□

Proof of Proposition 21 (sufficiency). The proof is very similar to that of Proposition 21, except that in the induction base we need to take into account that g may also be $\mathrm{L}_2(\cdot, 0)$ (or $\Gamma_2(\cdot, 1)$); this case is covered by Remark 2. Respectively, in the induction step, g can also be $\mathrm{L}_q(\cdot, 0)$. □

Example 1. When $g = \mathit{0}_q$, running the recursive definitions (30)–(33) with the initial conditions (27)–(28) yields, for every $j \in [q-1\rangle$, $x \in [q\rangle$, and $f \in \mathcal{F}_q \setminus \{\mathit{0}_q\}$:

$$u_j(x) = \begin{cases} 0 & \text{if } x \leq j \\ 1 & \text{if } x = j+1 \\ * & \text{if } x > j+1 \end{cases}$$

and

$$\vartheta_j(f) = \begin{cases} 0 & \text{if } f(j{+}1) = 0 \text{ and } f(x) = 1 \text{ for some } x \leq j \\ 1 & \text{if } f(j{+}1) = 1 \text{ and } f(x) = 0 \text{ for all } x \leq j \\ * & \text{otherwise} \end{cases} .$$

In words, $\boldsymbol{\vartheta}(f)$ is obtained from the truth table of f by changing all the 0's that precede the first 1 and all the 1's that succeed it into $*$'s, and then deleting the first entry (which corresponds to $x = 0$). Table IX shows the mappings $x \mapsto \boldsymbol{u}(x)$ and $f \mapsto \boldsymbol{\vartheta}(f)$ for $q = 4$. □

TABLE IX
MAPPINGS FOR $\mathcal{F}_4 \setminus \{\mathit{0}_4\}$ THAT ATTAIN $q = n+1 = 4$ UNDER SCENARIO $({*}{*})$.

x	$\boldsymbol{u}(x)$	$\boldsymbol{f}$	$\boldsymbol{\vartheta}(f)$	$\boldsymbol{f}$	$\boldsymbol{\vartheta}(f)$	$\boldsymbol{f}$	$\boldsymbol{\vartheta}(f)$
0	000	1000	000	0110	$1{*}0$	1101	${*}0{*}$
1	100	0100	100	1110	${*}{*}0$	0011	${*}1{*}$
2	${*}10$	1100	${*}00$	0001	${*}{*}1$	1011	$0{*}{*}$
3	${*}{*}1$	0010	${*}10$	1001	$00{*}$	0111	$1{*}{*}$
		1010	$0{*}0$	0101	$10{*}$	1111	${*}{*}{*}$

When $g \neq \mathit{0}_q$ we can assume that $g(x) = \mathrm{L}_q(x, t)$ for some $t \in [1 : q-1\rangle$ (or $t \in [q-1\rangle$ for Scenario $(\bullet{*})$). We run the recursive definitions (30)–(33) yet now with the initial conditions (22)–(25) when stated for $q = n+1 = t+2$. Table X shows the resulting mapping $x \mapsto \boldsymbol{u}(x)$ for $q = 8$ and $t = 3$. The entries $u_j(x)$ that correspond to $(x, j) \in [t+2\rangle \times [t+1\rangle$ are determined by the initial conditions and are the same as in Table VII, while the entries that correspond to $(x+t+1, j+t+1)$ for $(x, j) \in [q-t-1\rangle \times [q-t-2\rangle$ are the same as in Table IX.

TABLE X
MAPPING $x \mapsto \boldsymbol{u}(x)$ THAT ATTAINS $q = n+1$ FOR $q = 8$ AND $t = 3$.

x	$\boldsymbol{u}(x)$	
0	$1{*}{*}0$	${*}{*}{*}$
1	$01{*}{*}$	${*}{*}{*}$
2	${*}01{*}$	${*}{*}{*}$
3	${*}{*}01$	${*}{*}{*}$
4	0000	000
5	${*}{*}{*}{*}$	100
6	${*}{*}{*}{*}$	${*}10$
7	${*}{*}{*}{*}$	${*}{*}1$

Proof of Proposition 22. Suppose that $\mathcal{F}_q^\star$ is n-cell implementable with mappings $\boldsymbol{u} : [q\rangle \to \mathbb{B}_*^n$ and $\boldsymbol{\vartheta} : \mathcal{F}_q^\star \to \mathbb{B}_*^n$. Since $q \geq 3$, for any two distinct elements $x, x' \in [q\rangle$ there exists a function $f \in \mathcal{F}_q^\star$ such that $f(x) = 1$ yet $f(x') = 0$. Hence, the mapping $\boldsymbol{u}$ has to be injective and its set of images must form an antichain in $\mathbb{B}_*^n$ (see Section II).

For any element $t \in [q\rangle$ let t' be a particular element in $[q\rangle \setminus \{t\}$ (say, $t' = 1$ when $t = 0$, and $t' = 0$ otherwise). By the antichain property it follows that there is at least one index $\ell = \ell(t) \in [n\rangle$ such that $u_\ell(t) \not\preceq u_\ell(t')$ and, in particular, $u_\ell(t) \neq u_\ell(t')$, $u_\ell(t) \neq \bullet$, and $u_\ell(t') \neq *$. We define $b(t) \in \mathbb{B}$ by

$$b(t) = \begin{cases} u_\ell(t) & \text{if } u_\ell(t) \in \mathbb{B} \\ 1 & \text{if } u_\ell(t) = * \text{ and } u_\ell(t') \in \{0, \bullet\} \\ 0 & \text{if } u_\ell(t) = * \text{ and } u_\ell(t') = 1 \end{cases} .$$

Notice that $\mathrm{T}(u_\ell(t), b(t)) = 1$ yet $\mathrm{T}(u_\ell(t'), b(t)) = 0$. We also define $f_t : [q\rangle \to \mathbb{B}$ to be the function in $\mathcal{F}_q^\star$ which evaluates to 1 when $x \in \{t, t'\}$ and to 0 otherwise.

We now extend the mapping $\boldsymbol{\vartheta}$ to the domain $\mathcal{F}_q$ by

$$\boldsymbol{\vartheta}(\mathit{1}_q) = {*}{*}\ldots{*}$$

and, for every $t \in [q\rangle$ and $j \in [n\rangle$,

$$\vartheta_j(\mathrm{E}_q(\cdot,t)) = \begin{cases} \vartheta_j(f_t) & \text{if } j \neq \ell(t) \\ b(t) & \text{if } j = \ell(t) \end{cases} .$$

It can be readily verified that Eq. (2) holds for the added functions $\mathit{1}_q$ and $\mathrm{E}_q(\cdot,t)$, $t \in [q\rangle$, and, therefore, for the whole set $\mathcal{F}_q$. Hence, $\mathcal{F}_q$ is n-cell implementable. $\square$

ACKNOWLEDGMENT

I would like to express my thanks to Giacomo Pedretti from Hewlett Packard Labs. This work resulted from many stimulating discussions that I had with him during my visit at Labs. Thanks are also due to Luca Buoanno and Cat Graves.

REFERENCES

[1] I. Anderson, *Combinatorics of Finite Sets,* Dover Publications, Mineola, New York, 1987.

[2] M. Anthony, N. Biggs, *Computational Learning Theory,* Cambridge University Press, Cambridge, UK, 1992.

[3] A. Birkendorf, N. Klasner, C. Kuhlmann, H.U. Simon, “Structural results about exact learning with unspecified attribute values,” *J. Comput. Syst. Sci.,* 60 (2000), 258–277.

[4] A. Bremler-Barr, D. Hay, D. Hendler, “Layered interval codes for TCAM-based classification,” *Comput. Netw.,* 56 (2012), 3023–3039.

[5] W. Fraczak, W. Rytter, M. Yazdani, “Matching integer intervals by minimal sets of binary words with don’t cares,” *Proc. 19th Annual Symp. on Combinatorial Pattern Matching (CPM 2008),* Pisa, Italy, 2008, pp. 217–228.

[6] R. Govindaraj, S. Ghosh, “Design and analysis of STTRAM-based ternary content addressable memory cell,” *ACM J. Emerg. Technol. Comput. Syst.,* 13 (2017), Article 52, 1–22.

[7] G. Kasai, Y. Takarabe, K. Furumi, M. Yoneda, “200MH/200MSPS 3.2W at 1.5V Vdd, 9.4Mbits ternary CAM with new charge injection match detect circuits and bank selection scheme,” *Proc. IEEE 2003 Custom Integrated Circuits Conf.,* San Jose, CA, 2003, pp. 387–380.

[8] C. Li, C.E. Graves, X. Sheng, D. Miller, M. Foltin, G. Pedretti, J.P. Strachan, “Analog content-addressable memories with memristors,” *Nat. Commun.,* 11 (2020), #1638.

[9] D. Lubell, “A short proof of Sperner’s theorem,” *J. Comb. Theory,* 1 (1966), 299.

[10] F.J. MacWilliams, N.J.A. Sloane, *The Theory of Error-Correcting Codes,* North-Holland, Amsterdam, 1977.

[11] T. Natschläger, M. Schmitt, “Exact VC-dimension of Boolean monomials,” *Inf. Process. Lett.,* 59 (1996), 19–20.

[12] G. Pedretti, C.E. Graves, C. Li, S. Serebryakov, X. Sheng, M. Foltin, R. Mao, J.P. Strachan, “Tree-based machine learning performed in-memory with memristive analog CAM,” *Nat. Commun.,* 12 (2021), #5806.

[13] X. Yang, S. Sezer, J. McCanny, D. Burns, “A versatile content addressable memory architecture,” *Proc. 20th Anniversary IEEE Int’l SOC Conf.,* Hsinchu, Taiwan, 2007, pp. 215–218.