Thermal-Aware Channel with Multiple Wires

Chee, Yeow Meng; Etzion, Tuvi; Schouhamer Immink, Kees A. ; Nguyen, Tuan Thanh ; Vu, Van Khu ; Weber, Jos H.; Yaakobi, Eitan

DOI
10.1109/ISIT57864.2024.10619462

Publication date
2024

Document Version
Final published version

Published in
Proceedings of the 2024 IEEE International Symposium on Information Theory (ISIT)

Thermal-Aware Channel with Multiple Wires

Yeow Meng Chee*, Tuvi Etzion†, Kees A. Schouhamer Immink‡, Tuan Thanh Nguyen§, Van Khu Vu*, Jos H. Weber¶, and Eitan Yaakobi†

* Department of Industrial Systems Engineering and Management, National University of Singapore, Singapore
†Department of Computer Science, Technion-Israel Institute of Technology, Haifa, 3200003 Israel
‡Turing Machines Inc, Willemskade 15d, 3016 DK Rotterdam, The Netherlands
§ Science, Mathematics and Technology Cluster, Singapore University of Technology and Design, Singapore 487372
¶Department of Applied Mathematics, Delft University of Technology, 2628 CD Delft, The Netherlands

2024 IEEE International Symposium on Information Theory (ISIT) | 979-8-3503-8284-6/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISIT57864.2024.10619462

***Abstract*—The thermal-aware channel has been studied recently to control the temperature of some electronic devices for better performance and longer lifetime. In this work, we consider a thermal-aware channel model where multiple wires are available to the user. The user can use one wire or several wires to write an information word. Particularly, we study the two extreme cases. In the first case, only one wire is permitted for writing the information. The other extreme case is that we are allowed to write information on all the wires in parallel. In the first case, when we send a message through a wire that reaches the highest allowed temperature, we switch to another available wire. We determine the minimum number of wires required to send any arbitrary message. Given the number of wires, our second task is to determine the constrained codewords that can be sent through these wires. We compute the maximum information rate achieved and provide some constructions of codes satisfying these constraints. In the second case when all the wires are available for writing many, interesting questions arise and we briefly describe one of them and its solutions.**

I. Introduction

Power and heat dissipation have emerged as first-order design constraints for chips, whether targeted for battery-powered devices or for high-end systems. High temperatures have dramatic negative effects on bus performance. Power-aware design alone is insufficient to address the thermal challenges since it does not directly target the spatial and temporal behavior of the operating environment. For this reason, thermally aware approaches have emerged as one of the most important domains of research in chip design today. Numerous techniques have been proposed to reduce the overall power consumption of on-chip buses (see [1] which uses coding techniques and the references therein using non-coding techniques). More recent work on such problems can be found in [2]–[4]. All the non-coding techniques do not directly address peak temperature minimization. The coding techniques such as in [1], [2] consider the channel of multiple wires while the recent coding scheme [5] considers the channel of a single wire.

In this work, we consider the practical model in which the user wants to send information on multiple wires without being concerned about the possibility that one of the wires will overheat and cause information loss. The model that will be presented leads us from the single-wire model to the multi-wire model in two steps. First, the multiple-wire model will behave as a model to send one bit of information at a time. When a wire gets to its maximum temperature before being overheated, the information is sent to another available wire. Since in each time slot, only one bit of information is allowed to be sent, the model cannot be considered the same as the one in which it is permitted to send information in parallel on all the wires. Most of the paper will be devoted to this case. In the other extreme case, we are allowed to send information on all the wires in parallel. In this case, we ask the following natural question: what is the minimum number of required wires such that it will be able to send κ bits of information without having any overheated wires? This question will be briefly discussed and a comprehensive discussion is left for the full version of this paper.

II. Notations and Definitions

We start with a description of the thermal-aware channel with a single wire model, as was introduced in [5]. It is assumed that the temperature increases by $t_1 > 0$ after a current 'on' period and the temperature decreases by $t_0 > 0$, after an 'off' period. The maximum allowed temperature is $t_{\max}$ and the base temperature is $t_{\min}$. All temperature variables are non-negative real numbers. Set $M = t_{\max}/t_0$ and $k = t_1/t_0$.

Without loss of generality (w.l.o.g. in short), we assume that $t_{\min} = 0$, $t_0 = 1$, and hence, $M = t_{\max}$ and $k = t_1$.

Definition 1. Given $k > 0$ and a binary sequence $\boldsymbol{x} \in \{0,1\}^n$, the *k-temperature-sequence* of $\boldsymbol{x}$, denoted by $\Phi_k(\boldsymbol{x})$, is a sequence $\boldsymbol{s} = \Phi_k(\boldsymbol{x}) = s_0 s_1 s_2 \ldots s_n$, where $s_0 = 0$ and for all $1 \leqslant i \leqslant n$, $s_i = s_{i-1} + k$ if $x_i = 1$ and $s_i = \max\{s_{i-1} - 1, 0\}$ if $x_i = 0$.

Definition 2. Let M, k be some positive real numbers.

- A sequence $\boldsymbol{x} \in \{0,1\}^n$ will be called an (M,k)-*thermal-aware sequence* (TA sequence in short) if its *k-temperature-sequence* $s = \Phi_k(\boldsymbol{x}) = s_0 s_1 \ldots s_n$, satisfies that $0 \leqslant s_i \leqslant M$ for all $0 \leqslant i \leqslant n$.
- The set of all (M,k)-TA sequences of length n is called the *maximal* (M,k)-*thermal-aware code* (or TA code), and is denoted by $\mathcal{A}(M,k,n)$.
- The channel that only accepts (M,k)-TA sequences is called the (M,k)-*thermal-aware channel* (or TA channel). The *capacity* of the (M,k)-TA channel, is defined as:

$$\mathsf{cap}_{\mathsf{TA}}(M,k) = \limsup_{n\to\infty} \frac{\log_2 |\mathcal{A}(M,k,n)|}{n}.$$

In [5], the exact capacity of the channel with certain parameters was derived and several bounds on the capacity in various cases were presented. In the next subsection, we formally define the *thermal-aware channel with multiple wires*.

A. Thermal-aware Channel with Multiple Wires

In this model, we assume that there is more than one wire, but the wires are connected in a way that we can send a pulse only on one of them. When the chip reaches its maximum allowed temperature in some of the wires, the pulses can be applied only to one of the other available wires.

Given positive integers a, b, let $[a;b]$ denote the set $\{a, a+1, a+2, \ldots, b\}$ and let $[n]$ denote the set $\{1, 2, \ldots, n\}$. Given a binary sequence $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \{0,1\}^n$ and a set $I \subseteq \{1, 2, \ldots, n\}$, the *complete-projection* of $\boldsymbol{x}$ in I, denoted by $\mathrm{P}_I(\boldsymbol{x})$, is the binary sequence $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ where $y_i = x_i$ if $i \in I$, and $y_i = 0$ otherwise. For example, consider $\boldsymbol{x} = 100101$ and $I_1 = \{1, 3, 4, 5\}$ and $I_2 = \{2, 6\}$. We then have $\mathrm{P}_{I_1}(\boldsymbol{x}) = 100100$ and $\mathrm{P}_{I_2}(\boldsymbol{x}) = 000001$. The thermal-aware multiple-wire model is formally defined as follows.

Definition 3. Let M, k be some positive real numbers and $m \geqslant 1$ denote the number of wires.

- A sequence $\boldsymbol{x} \in \{0,1\}^n$ is called an $(m; M, k)$*-TA sequence* if there exists a partition of all indices in $\boldsymbol{x}$ into m mutually disjoint subsets $I_1, I_2, \ldots, I_m$, such that $\bigcup_{i\in[m]} I_i = [n]$ and for all $1 \leqslant j \leqslant m$, the complete-projection of $\boldsymbol{x}$ in I_j, $\mathrm{P}_{I_j}(\boldsymbol{x})$, is an (M, k)-TA sequence.
- The set of all $(m; M, k)$-TA sequences of length n is called the *maximal* $(m; M, k)$*-TA code*, and is denoted by $\mathcal{A}(m; M, k, n)$.
- The channel that only accepts $(m; M, k)$-TA sequences is called the $(m; M, k)$*-TA channel*. The *capacity* of this channel is the maximum achievable asymptotic rate, i.e.,

$$\mathsf{cap}_{\mathsf{TA}}(m; M, k) = \limsup_{n\to\infty} \frac{\log_2 |\mathcal{A}(m; M, k, n)|}{n}.$$

  where the division is only by n since we are forced to write only on one wire by the channel constraint.

Remark 1. We observe that if $\boldsymbol{x}$ is an $(m; M, k)$ TA-sequence, then there exists an efficient encoding scheme over m wires, that the electric pulses are applied on the jth wire on those indices $i \in I_j$, and the maximum temperature over each of the m wires is at most M. On the other hand, to decode the transmitted signals, it is not necessary to require every $(m; M, k)$-sequence to share the same m partitions of indices $I_1, I_2, \ldots, I_m$. Since $I_{j_1} \cap I_{j_2} = \emptyset$ for all $j_1 \neq j_2$, a transmitted signal $\boldsymbol{x}$ can be decoded uniquely by implying the bitwise OR over all the sequences. For example, if $\mathrm{P}_{I_1}(\boldsymbol{x}) = 110001$ and $\mathrm{P}_{I_2}(\boldsymbol{x}) = 000110$ then $\boldsymbol{x} = 110001$ **OR** $000110 = 110111$.

B. Summary of the Results in the Paper

We outline our sections and the main contributions of this work as follows.

Determining the minimum number of wires. To minimize the cost and power supply, our first aim is to determine $\nu(M, k, n)$, which is the minimum number of wires necessary to transmit *any* binary sequence of length n, i.e., given $M, k, n > 0$,

$$\nu(M,k,n) \triangleq \min\left\{m : \boldsymbol{x} \text{ is an } (m; M, k, n) \text{ TA-sequence } \forall \boldsymbol{x} \in \{0,1\}^n\right\}.$$

It was shown in [5] that $\nu(M, k, n) \leqslant \lceil k \rceil + 1$ for all M, n. In Section III, we provide an explicit formula to determine exactly the value of $\nu(M, k, n)$, which depends on M, k, n.

Computing the channel capacity $\mathsf{cap}_{\mathsf{TA}}(m; M, k)$. In Section IV, we modify the method in [5] (which computes the capacity of the channel with one wire $\mathsf{cap}_{\mathsf{TA}}(M, k)$) to compute the capacity of the channel with m wires $\mathsf{cap}_{\mathsf{TA}}(m; M, k)$. Such a method applies to any set of channel parameters, however, the complexity grows exponentially with M.

Providing efficient constructions of TA codes with two wires. To overcome the challenge of large M, in Section V, we consider the TA channel with two wires and provide some explicit constructions of TA codes associated with low-complexity encoding/decoding algorithms.

Answering some questions regarding the minimum number of wires required to send κ bits of information in parallel. This of course depends on the knowledge of both the transmitter and the receiver regarding the current temperature of the wires or more precisely which wires might be overheated if they have an 'on' pulse in the current round of communication (see Section VI).

III. Determining the Minimum Number of Wires

Recall that a binary sequence $\boldsymbol{x}$ is an $(m; M, k)$ TA sequence, if there exists a partition of all indices in $\boldsymbol{x}$ into m mutually disjoint subsets $I_1, I_2, \ldots, I_m$, such that $\bigcup_{i\in[m]} I_i = [n]$ and for all $1 \leqslant j \leqslant m$, the complete-projection of $\boldsymbol{x}$ in I_j, $\mathrm{P}_{I_j}(\boldsymbol{x})$, is an (M, k)-TA sequence. In other words, to verify if a binary sequence is an $(m; M, k)$-TA sequence, one may need to check every possible m mutually disjoint subsets $I_1, I_2, \ldots, I_m$, which is an extremely time-consuming process. Hence, a natural question is the following: Is there a pre-determined set of m mutually disjoint subsets $I_1^*, I_2^*, \ldots, I_m^*$ such that a binary sequence $\boldsymbol{x}$ is an $(m; M, k)$-TA sequence if and only if for all $1 \leqslant j \leqslant m$, the complete-projection of $\boldsymbol{x}$ in I_j^*, $\mathrm{P}_{I_j^*}(\boldsymbol{x})$, is an (M, k)-TA sequence? We observe that it is sufficient to only consider the partition of indices i where $x_i = 1$. Indeed, for an entry $x_i = 0$, it implies that $\mathrm{P}_{I_j^*}(\boldsymbol{x})_i = 0$ for all $1 \leqslant j \leqslant m$.

Lemma 1. *Consider a binary sequence $\boldsymbol{x} \in \{0,1\}^n$. Let $\mathrm{P}(\boldsymbol{x}) = \{i : x_i = 1\}$. Let $I_1^*, I_2^*, \ldots, I_m^*$ be m mutually disjoint subsets such that $\bigcup_{j\in[m]} I_j^* = [n]$ and for $1 \leqslant s \leqslant |\mathrm{P}(\boldsymbol{x})|$, $1 \leqslant j \leqslant m$, the s-th element in $\mathrm{P}(\boldsymbol{x})$ is in I_j^* if $s \equiv j \pmod m$. We then have that $\boldsymbol{x}$ is an $(m; M, 1, k)$-TA sequence if and only if $\mathrm{P}_{I_j^*}(\boldsymbol{x})$ is an (M, k)-TA sequence for all $1 \leqslant j \leqslant m$.*

Proof. We observe that the m subsets $I_1^*, I_2^*, \ldots, I_m^*$ are constructed in such a way that at every step the electric pulse is applied on a wire with the minimum temperature among m wires.

For $1 \leqslant j \leqslant m$, we use $\vartheta_j(t)$ to denote the difference between the maximum temperature allowed and the current temperature on the j-th wire at time t, and we set $\vartheta(t) = \sum_{j=1}^m \vartheta_j(t)$. Suppose that at time t, we have a set S of wires with the same minimum temperature $h \geqslant 0$ while the set $S' = [m] \setminus S$ includes those wires with strictly higher temperatures. We have two following cases:

- If $h > 0$, we observe that applying the electric pulse in any wire results in the same value of $\vartheta(t+1) = \vartheta(t) + (m-1) - k$ since there are exactly $(m-1)$ wires that are cooling down by 1 while one wire is heating up by k.
- If $h = 0$, we observe that applying the electric pulse in a wire in S would result in a higher value of $\vartheta(t+1)$. If we apply the electric pulse in a wire in S, we have $\vartheta(t+1) = \vartheta(t) + |S'| - k = v_1$. On the other hand, if we apply the electric pulse in any wire in S', we have $\vartheta(t+1) = \vartheta(t) + (|S'| - 1) - k = v_2 = v_1 - 1 < v_1$.

In conclusion, if the electric pulse is applied on a wire with the minimum temperature among m wires, we have a higher value of $\vartheta(t)$, consequently giving us more options to apply the electric pulse in time $t+1$. In other words, if the complete-projections of $\boldsymbol{x}$ in the m mutually disjoint subsets $I_1^*, I_2^*, \ldots, I_m^*$ are not all (M,k)-TA sequences then for any arbitrary partition with m mutually disjoint subsets $I_1, I_2, \ldots, I_m$, the complete-projections of $\boldsymbol{x}$ in $I_1, I_2, \ldots, I_m$ are not all (M,k)-TA sequences. ■

Example 1. Consider $M = 6$, $k = 3$, $m = 3$ and a sequence $\boldsymbol{x} = 111111111111 \in \{0,1\}^{12}$. We first consider a partition $I_1 = \{1,2,7,8\}$, $I_2 = \{3,4,9,10\}$, $I_3 = \{5,6,11,12\}$, and

$$\mathrm{P}_{I_1}(\boldsymbol{x}) = 110000110000,$$
$$\mathrm{P}_{I_2}(\boldsymbol{x}) = 001100001100,$$
$$\mathrm{P}_{I_3}(\boldsymbol{x}) = 000011000011.$$

We verify that all $\mathrm{P}_{I_1}(\boldsymbol{x})$, $\mathrm{P}_{I_2}(\boldsymbol{x})$, and $\mathrm{P}_{I_3}(\boldsymbol{x})$ are not $(6,3)$-TA sequences. On the other hand, according to Lemma 1, we can set $I_1^* = \{1,4,7,10\}$, $I_2^* = \{2,5,8,11\}$ and $I_3^* = \{3,6,9,12\}$. We then obtain

$$\mathrm{P}_{I_1^*}(\boldsymbol{x}) = 100100100100,$$
$$\mathrm{P}_{I_2^*}(\boldsymbol{x}) = 010010010010,$$
$$\mathrm{P}_{I_3^*}(\boldsymbol{x}) = 001001001001.$$

We verify that all $\mathrm{P}_{I_1^*}(\boldsymbol{x})$, $\mathrm{P}_{I_2^*}(\boldsymbol{x})$, and $\mathrm{P}_{I_3^*}(\boldsymbol{x})$ are $(6,3)$-TA sequences. Thus, $\boldsymbol{x}$ is a $(3;6,3)$-TA sequence.

To determine the value of $\nu(M,k,n)$, i.e., the minimum number of wires necessary to transmit every binary sequence $\boldsymbol{x} \in \{0,1\}^n$, one may consider the following equivalent problem. Given $M, k > 0$ and $m > 0$ as the number of wires, determine the largest integer n such that every binary sequence of length n is an $(m;M,k,n)$-TA sequence. Formally, we define:

$$\rho(m;M,k) \triangleq$$
$$\max\Big\{n : \boldsymbol{x} \text{ is } (m;M,k,n)\text{-TA sequence for all } \boldsymbol{x} \in \{0,1\}^n\Big\}.$$

The following result is then immediate.

Lemma 2. *Given $M \geqslant k > 0$, we have*

(i) $\nu(M,k,n) = m$ *if and only if*

$$\rho(m-1;M,k) < n \leqslant \rho(m;M,k).$$

(ii) $\nu(M,k,n) \leqslant \lceil k \rceil + 1$. *Equivalently, for $m \geqslant \lceil k \rceil + 1$, we have $\rho(m;M,k) = \infty$.*

(iii) For $n \leqslant \lfloor M/k \rfloor$, we have $\nu(M,k,n) = 1$. Equivalently, for $m = 1$, we have $\rho(1;M,k) = \lfloor M/k \rfloor$.

Theorem 1. *Given $M \geqslant k > 1$ and $m \leqslant \lceil k \rceil$, we have*

$$\rho(m;M,k) = m \left\lfloor \frac{M-m+1}{k-m+1} \right\rfloor.$$

Proof. We consider the worst-case scenario when the transmitted signal is the all ones vector, i.e., $\boldsymbol{x} = 1^n$. For simplicity, suppose that $n = mn_0$. According to Lemma 1, $\boldsymbol{x}$ is an $(m;M,k,n)$-TA sequence if and only if $\mathrm{P}_{I_i^*}(\boldsymbol{x})$ is an (M,k)-TA sequence for $1 \leqslant i \leqslant m$, where $\mathrm{P}_{I_i^*}(\boldsymbol{x}) = (0^{i-1}10^{m-i})^{n/m}$. By computing the temperature vector of $\mathrm{P}_{I_1^*}(\boldsymbol{x})$, we then have the maximum temperature is at the $(n-m+1)$-th index, which must be bounded above by M, i.e.,

$$\Big(k-m+1\Big)\frac{n}{m} + m - 1 \leqslant M,$$
$$\frac{n}{m} \leqslant \frac{M-m+1}{k-m+1},$$
$$n \leqslant m \left\lfloor \frac{M-m+1}{k-m+1} \right\rfloor. \qquad ■$$

From Lemma 2, one can determine the value of $\nu(M,k,n)$ according to the value of $\rho(m;M,k)$ and $\rho(m-1;M,k)$, where the explicit formula is given in Theorem 1.

Example 2. Given $M = 50, k = 3$. According to Theorem 1, we compute $\rho(2;50,3) = 48$, $\rho(3;50,3) = 144$, and $\rho(1;50,3) = 16$. Consequently, we claim that:

- $\nu(50,3,n) = 1$ for all $n \leqslant 16$,
- $\nu(50,3,n) = 2$ for all $16 < n \leqslant 48$,
- $\nu(50,3,n) = 3$ for all $48 < n \leqslant 144$, and
- $\nu(50,3,n) = 4$ for all $n > 144$.

We formally state the result of $\nu(M,k,n)$ as follows.

Corollary 1. *Given $M \geqslant k \geqslant 1$. We have $\nu(M,k,n) = 1$ if and only if $n \leqslant \lfloor M/k \rfloor$. In addition, for $2 \leqslant i \leqslant \lceil k \rceil$, we have $\nu(M,k,n) = i$ if and only if*

$$(i-1)\left\lfloor \frac{M-i+2}{k-i+2} \right\rfloor < n \leqslant i \left\lfloor \frac{M-i+1}{k-i+1} \right\rfloor.$$

Lastly, $\nu(M,k,n) = \lceil k \rceil + 1$ for all $n > \lceil k \rceil \left\lfloor \frac{M-\lceil k \rceil+1}{k-\lceil k \rceil+1} \right\rfloor$.

IV. Capacity of the Multiple-Wire TA Channel

In this section, we present a method to compute the capacity of the TA channel with m wires, $\mathsf{cap}_{\mathsf{TA}}(m;M,k)$ for $m > 1$. Recall that for $m \geqslant k+1$, we have $|\mathcal{A}(m;M,k,n)| = 2^n$, and hence $\mathsf{cap}_{\mathsf{TA}}(m;M,k) \equiv 1$. We now consider the channel with two wires and $k \geq 2$, and for simplicity, we further assume that M,k are integers. It is known that not all sequences can be sent through the channel of two wires, for example, the all-ones sequence can not be sent through the channel. Hence, we need to determine which sequences can be sent through the channel, and thus, we can compute the capacity of the channel. We first build a finite state machine of the channel which is a labelled graph $G_{2;M,k}$ of $(M+1)^2$ vertices, denoted by

$$S_2 = \Big\{(a,b) : a,b \in \{0,1,2,\ldots,M\}\Big\}.$$

Each vertex (a,b) of the graph is the state (a,b) that the temperature of the first wire is a and the temperature of the second wire is b. There is a directed edge from vertex (a_1, b_1) to vertex (a_2, b_2) if a transition from state (a_1, b_1) to state (a_2, b_2) is possible.

The adjacency matrix $D_{2;M,k}$ of size $(M+1)^2 \times (M+1)^2$ of the graph $G_{2;M,k}$ is defined such that $d_{2;M,k}((a_1, b_1), (a_2, b_2)) = 1$ if there is a directed edge from vertex (a_1, b_1) to vertex (a_2, b_2) and $d_{2;M,k}((a_1, b_1), (a_2, b_2)) = 0$ otherwise. To describe $G_{2;M,k}$ in details, we describe its adjacency matrix $D_{2;M,k}$ of size $(M+1)^2 \times (M+1)^2$ as follows. For $1 \leqslant k \leqslant M$, $0 \leqslant a, b \leqslant M$,

$$d_{2;M,k}((0,0),(0,0)) = 1;$$
$$d_{2;M,k}((a,b),(\max\{a-1,0\}, b+k)) = 1 \text{ if } b \leqslant M-k;$$
$$d_{2;M,k}((a,b),(a+k, \max\{b-1,0\})) = 1 \text{ if } a \leqslant M-k;$$
$$d_{2;M,k}((a,b),(\max\{a-1,0\}, \max\{b-1,0\})) = 1;$$
$$d_{2;M,k}((a_1,b_1),(a_2,b_2)) = 0, \text{ otherwise.}$$

We see that the graph $G_{2;M,k}$ is the finite state machine of the $(2; M, k)$ thermal-aware channel, that is, for each path of length n in the graph, there is a sequence of length n that can be sent through the channel. For all valid sequences in the channel, there is a corresponding path in the graph $G_{2;M,k}$. However, this correspondence is not a one-to-one mapping. For example, in case $M = 4, k = 2$, from the initial state $(0, 0)$, when we send the bit '1' the following state would be either $(0, 2)$ or $(2, 0)$. The graph $G_{2;M,k}$ is known as non-deterministic finite automaton (NFA).

To compute the number of valid sequences in the channel and to compute the capacity of the channel, we need to build a deterministic finite automaton (DFA) where there is a one-to-one mapping between a valid sequence in the channel and a path in the graph. It is known that each NFA can be converted into an equivalent DFA [6]. Once we have an equivalent DFA, we can compute the capacity of the channel by computing the base-2 logarithm of the largest real eigenvalue of the adjacency matrix of the DFA. Using this technique, we compute the capacity of the TA channel for various parameters M, k, and tabulate the numerical results in Table I. We also present the following example to illustrate the process of computing the capacity for the case $M = 2$ and $k = 2$.

Example 3. For $M = 2, k = 2$, the adjacency matrix $D_{2;2,2}$ of the NFA $G_{2;2,2}$ is computed as follows. Here, the states are listed in lexicographic order, i.e., $(0,0) \to 1, (0,1) \to 2, (0,2) \to 3, (1,0) \to 4, (1,1) \to 5, (1,2) \to 6, (2,0) \to 7, (2,1) \to 8$, and $(2,2) \to 9$. $S = \{1,2,3,4,5,6,7,8,9\}$.

$$D_{2;2,2} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

We then convert the non-deterministic finite automaton to a deterministic finite automaton with a corresponding list of states

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$M=2$	0.0						
$M=3$	0.8791	0.0					
$M=4$	0.9698	0.7776	0.0				
$M=5$	0.9904	0.8571	0.6979	0.0			
$M=6$	0.9967	0.9190	0.7623	0.6346	0.0		
$M=7$	0.9988	0.9436	0.8196	0.6870	0.5832	0.0	
$M=8$	0.9995	0.9598	0.8623	0.7357	0.6265	0.5405	0.0
$M=9$	0.9998	0.9695	0.8841	0.7768	0.6676	0.5769	0.5045

TABLE I: The capacity of the TA channel with two wires.

as follows $S' = \{\emptyset, 1, 2, 3, 4, \{2,4\}, 6, 7, \{3,7\}, 8, \{6,8\}\}$ and a corresponding adjacency matrix $D'_{2;2,2}$ as follows.

$$D'_{2;2,2} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The largest real eigenvalue of $D'_{2;2,2}$ is $\lambda = 1.8393$, and thus, the channel capacity is $\mathsf{cap}_{\mathsf{TA}}(2; 2, 2) = \log_2 1.8393 = 0.8791$.

This method can be extended to the case of having m wires, that is, given all three parameters m, M, k, we can compute exactly the capacity of m-wire thermal-aware channel, $\mathsf{cap}_{\mathsf{TA}}(m; M, k)$. The adjacency matrix $D_{m;M,k}$ of the NFA can be computed in the same way as the case $m = 2$ above. However, the algorithm to convert from NFA to DFA is not efficient as the computational complexity grows exponentially with M. To overcome the challenge for large M, in the next section, we provide some explicit constructions of TA codes associated with low-complexity encoding/decoding algorithms.

V. CONSTRUCTIONS OF TA CODES WITH TWO WIRES

A set $\mathcal{C} \subseteq \{0,1\}^n$ of $(m; M, k)$-TA sequences of length n is called an $(m; M, k, n)$-*TA code*. In this section, we consider the case $m = 2$ and provide some explicit constructions of $(2; M, k, n)$-TA codes for arbitrary given M, k, and n. We defer the study of $(m; M, k, n)$-TA codes for $m > 2$ for future work.

Recall that in Example 3, we compute $\mathsf{cap}_{\mathsf{TA}}(2; 2, 2) = \log_2 1.82993 = 0.87179$. The following construction provides a code whose asymptotic rate is $\log_2 3/2 \approx 0.7925$.

For $\Sigma_3 = \{0, 1, 2\}$, we define a one-to-one mapping $\Phi : \Sigma_3 \to \Sigma_2^2$ as follows:

$$\Phi(0) = (0,0), \Phi(1) = (0,1), \text{and } \Phi(2) = (1,0).$$

Consequently, given an arbitrary ternary sequence $\boldsymbol{x} = (x_1, x_2 \ldots, x_n) \in \Sigma_3^n$, we obtain the corresponding sequence $\boldsymbol{y} = \Phi(\boldsymbol{x}) = (y_1, \ldots, y_{2n}) \in \Sigma_2^{2n}$, where $\Phi(x_i) = (y_{2i-1}, y_{2i})$ for all $1 \leq i \leq n$.

Theorem 2. *Given $m = 2$, $k = 2$, $M \geqslant 2$, and $n > 0$ we set*

$$\mathcal{C} = \left\{ \Phi(\boldsymbol{x}) : \boldsymbol{x} \in \Sigma_3^n \right\}.$$

We then have that $\mathcal{C}$ is a $(2; M, 2, 2n)$-TA code of length $(2n)$. The size of $\mathcal{C}$ is $|\mathcal{C}| = 3^n$ with rate $\mathbf{r}_{\mathcal{C}} = \log_2 3/2 \approx 0.7925$.

Proof. It is sufficient to show that for arbitrary $\boldsymbol{c} \in \mathcal{C}$, $\boldsymbol{c}$ is a $(2; M, 2, 2n)$-TA sequence. We set $I_1 = \{4i+1, 4i+2 : i \geqslant 0\} \cap [2n]$ and $I_2 = \{4i, 4i-1 : i \geqslant 1\} \cap [2n]$. Suppose that $\boldsymbol{c} = \Phi(\boldsymbol{x})$ for some sequence $\boldsymbol{x} \in \Sigma_3^n$. We have two complete-projection sequences as

$$\mathrm{P}_{I_1}(\boldsymbol{c}) = c_1 c_2 0 0 c_5 c_6 0 0 \ldots,$$
$$\mathrm{P}_{I_2}(\boldsymbol{c}) = 0 0 c_3 c_4 0 0 c_7 c_8 \ldots.$$

We observe that $c_{2i-1}c_{2i} \neq 11$ for all $i \geqslant 1$. Therefore, the maximum temperature on both $\mathrm{P}_{I_1}(\boldsymbol{c})$ and $\mathrm{P}_{I_2}(\boldsymbol{c})$ is at most $2 \leqslant M$. Thus, $\boldsymbol{c}$ is a $(2; M, 2, 2n)$-TA sequence. ■

One may extend the idea of Theorem 2 to construct codes of larger M. For example, when $M \geqslant 4$, one may define the following one-to-one mapping $\Phi' : \Sigma_6 \to \Sigma_2^3$:

$$\Phi'(0) = (0,0,0), \Phi'(1) = (0,0,1), \Phi'(2) = (0,1,0),$$
$$\Phi'(3) = (1,0,0), \Phi'(4) = (1,0,1), \Phi'(5) = (1,1,0).$$

Similarly, we obtain a set $\mathcal{C}' = \left\{ \Phi'(\boldsymbol{x}) : \boldsymbol{x} \in \Sigma_6^n \right\}$, where $\mathcal{C}'$ is a $(2; M, 2, 3n)$-TA code of length $(3n)$. The size of $\mathcal{C}$ is $|\mathcal{C}| = 6^n$ and the rate is $\mathbf{r}_{\mathcal{C}'} = \log_2 6/2 \approx 0.8617$.

We now provide a construction for general $M, k, n > 0$. Given two sequences $\boldsymbol{x} = (x_1, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, \ldots, y_m)$, we use $\boldsymbol{xy}$ to denote the concatenation of these two sequences, i.e., $\boldsymbol{xy} = (x_1, \ldots, x_n, y_1, \ldots, y_m)$.

Theorem 3. *Let $\mathcal{C}_1$ be an $(1; M, k, M)$-TA code of length M. Given $n = aM$ for some positive integer a, we set*

$$\mathcal{C}_2 = \left\{ \boldsymbol{c} \in \{0,1\}^n : \boldsymbol{c} = \boldsymbol{x}_1 \boldsymbol{x}_2 \cdots \boldsymbol{x}_a : \boldsymbol{x}_i \in \mathcal{C}_1 \text{ for } 1 \leqslant i \leqslant a \right\}.$$

We then have that $\mathcal{C}_2$ is a $(2; M, k, n)$-TA code with rate $\log_2 |\mathcal{C}_1|/M$.

Proof. For each codeword $\boldsymbol{c} \in \mathcal{C}_2$, suppose that $\boldsymbol{c} = \boldsymbol{x}_1 \boldsymbol{x}_2 \cdots \boldsymbol{x}_a$. We write alternatively every M bits on the two wires. That is, we write $\boldsymbol{x}_1$ in the first wire and switch to write $\boldsymbol{x}_2$ in the second wire, then switch to the first wire to write $\boldsymbol{x}_3$, and so on. It is easy to verify that $\mathrm{P}_{I_1}(\boldsymbol{c})$ and $\mathrm{P}_{I_2}(\boldsymbol{c})$ are both $(M, 1, k)$-TA sequences. Hence, $\boldsymbol{c}$ is a $(2; M, k, n)$-TA sequence. We conclude that $\mathcal{C}_2$ is a $(2; M, k, n)$-TA code. ■

We now discuss the rate of $\mathcal{C}_2$. For example, when $M = 4$ and $k = 2$, we can exhaustively search for the optimal $(1; 4, 2, 4)$-TA code $\mathcal{C}_1$ of length 4, where

$$\mathcal{C}_1 = \Big\{ 0000, 0001, 0010, 0100, 1000,$$
$$0011, 0101, 0110, 1001, 1010, 1100 \Big\}.$$

The rate of the code $\mathcal{C}_2$ is then $\log_2 11/4 \approx 0.8648$. Similarly, when $M = 6, k = 2$, we obtain a code of rate 0.9287.

VI. Transmitting κ bits of information

In this section, we introduce the thermal-aware channel with multiple wires on which the transmitter should send information. The transmitter is submitting words of length κ that are the information words. To overcome the thermal constraints of the channel, the words of length κ are encoded onto m wires. There are two ways to encode the information for the channel with one laser diode (or a wire). The first way is to consider the channel operating on an 'on' and an 'off' status. When the laser is in an 'on' status it is considered as a '1' and the temperature is increasing. When the channel is in an 'off' status it is considered as a '0' and the temperature is decreasing. The second way is to use ***differential encoding*** and ***differential decoding***. In this method, the data to be communicated through the m wires is the difference between the current transmission and the previous one. This method has been widely used in thermal-aware channels, digital communication, and magnetic recording, among other applications. When there is a change in the information on a wire (either from '0' to '1' or from '1' to '0') the temperature increases and a '1' is transmitted. If the wire does not change its status during the transmission, then a '0' is transmitted on it. There are some differences between the two models. To make it clear, we summarize that differential encoding and differential decoding imply that to send the m-bit codeword, we usually have to know the current m bits word $\boldsymbol{v}$ written on the m wires and the m-bit codeword $\boldsymbol{c}$. The bit-by-bit addition of $\boldsymbol{v}$ and $\boldsymbol{c}$, i.e., $\boldsymbol{v} + \boldsymbol{c}$ is the new word written on the m wires. In this method the temperature on the wire is increased when the information on the wire is changed, i.e., the associated bit of the codeword is a *one* and the temperature is decreased when the associated bit of the codeword is a *zero*. It is clear that in this model either we know where exactly the hot wires are or we know where we can write in advance.

Theorem 4. *If there are no two consecutive 'on' pulses, then $m = 3\ell + 1$ wires are sufficient to encode $\kappa = 2\ell$ information bits when the transmitter and the receiver know which wires were 'on' in the previous transmission.*

Proof. At each round, there will be at most ℓ hottest wires. There are $2\ell + 1$ wires which are cool. Since

$$\sum_{i=0}^{\ell} \binom{2\ell+1}{i} = 2^{2\ell},$$

it follows that it is possible to encode $\kappa = 2\ell$ bits of information on these $2\ell + 1$ wires using at most ℓ wires that will be with 'on' pulse in the next round. ■

Corollary 2. *If there are no two consecutive 'on' pulses, then $m = 3\ell + 3$ wires are sufficient to encode $\kappa = 2\ell + 1$ information bits when the transmitter and the receiver know which wires were 'on' in the previous transmission.*

Theorem 5. *If there are no two consecutive 'on' pulses, then for κ large enough $m = 1.441\kappa$ wires are sufficient to encode κ information bits when the transmitter and the receiver know which wires were 'on' in the previous transmission.*

REFERENCES

[1] Y. M. Chee, T. Etzion, H. M. Kiah, and A. Vardy, "Cooling Codes: Thermal-Management Coding for High-Performance Interconnects," *IEEE Transactions on Information Theory,* vol. 64, no. 4, pp. 3062-3085, April 2018, doi: 10.1109/TIT.2017.2771245.

[2] Y. M. Chee, T. Etzion, H. M. Kiah, A. Vardy, and H. Wei, "Low-power cooling codes with efficient encoding and decoding, " *IEEE Transactions on Information Theory,* vol. 66, no. 8, pp. 4804–4818, 2020.

[3] U. Niesen and S. Kudekar, "Joint crosstalk-avoidance and error-correction coding for parallel data buses, " *IEEE Transactions on Information Theory,* vol. 65, no. 3, pp. 1626–1638, 2019.

[4] L. Valentini and M. Chiani, "Analysis and design of energy-efficient bus encoding schemes," *arxiv:2303.06409v1*.

[5] Y. M. Chee, T. Etzion, K. A. S. Immink, T. T. Nguyen, V. K. Vu, J. H. Weber, and E. Yaakobi, "Thermal-aware channel capacity," *Proc. Int. Symp. Inform. Theory*, 2023.

[6] J. Martin, *Introduction to Languages and the Theory of Computation*, McGraw Hill, 4th ed., 2010.