On the Problem of Finding "Sets Ensuring Linearly Independent Transversals" (SELIT), and its Application to Network Coding

Hirah Malik, Cédric Adjih, Michel Kieffer, Claudio Weidmann

On the Problem of Finding "Sets Ensuring Linearly Independent Transversals" (SELIT), and its Application to Network Coding

Hirah Malik, Cedric Adjih
Inria, Saclay
91120 Palaiseau, France
<name> . <surname>@inria.fr

Michel Kieffer
Université Paris-Saclay, CNRS
CentraleSupelec, L2S
91192 Gif-sur-Yvette, France
michel.kieffer@l2s.centralesupelec.fr

Claudio Weidmann
ETIS UMR8051, CY University,
ENSEA, CNRS
95000 Cergy, France
claudio.weidmann@ensea.fr

***Abstract*—This paper introduces a new formal mathematical problem initially motivated by an application of Network Coding (NC) to Information Centric Networks (ICN). It is of more limited scope but is remotely inspired by the well-known *index coding* problem. It is presented as follows: "given a vector space, can one construct several subsets of vectors, such that when drawing arbitrarily one vector from each subset, the selected vectors would be always linearly independent?". Answering this question is a step to construct an ICN efficient scheme with NC. We prove that our previously introduced construction is the only possible solution for a large family of constructions. This is an important result by itself. It also implies that any alternate solutions are outside this family and we propose one example.**

I. Introduction

The Efficient delivery of source content from some servers to one or several clients is one of the common network operations. In this paper, we are motivated by the use of Information Centric Networking [1] for such content delivery. IP protocol is not used; instead, clients send *Interest* packets to the network (not to a specific server); the Interests find their way to one or several servers with content; the content is then returned in *Data* packets on the reverse path. Additionally, content can be cached inside the network to potentially satisfy future requests.

Our starting point is the inefficiency illustrated by the scenario in Fig. 1, with basic ICN. Given a client and two servers, the client sends one request (Interest) labeled I_1 to the network to retrieve one unit of source content. Notice that the Interest I_1 is forwarded on two paths. The servers S_1 and S_2 reply with content $Q_1 = Q'_1 = P_1$. Nevertheless, only the first Data reply Q_1 is useful to the client, and Q'_1 is redundant.

Inefficiency occurs because 1) Interest I_1 is forwarded on more than one path and 2) because the Data packets Q_1 and Q'_1 carry identical content. The rationale for forwarding Interest I_1 on multiple paths is a potential throughput increase. The real issue is that Q_1 is equal to Q'_1: Q'_1 might have returned more useful information.

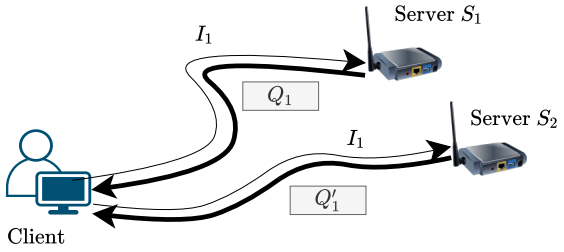

Fig. 1. Can ICN make efficient use of multiple paths?

One interesting solution is to adopt network coding(NC) [2], to improve the performance of ICN, as proposed in [3]. There is currently ongoing work in standardization on combining NC and ICN [4]. (Linear) network coding consists of viewing all packets as vectors of elements of a finite field and performing algebraic operations on them, such as linear combinations.

In our case, if the replies Q_1 and Q'_1 were coded packets, it could be possible for the servers to return, for example, $Q_1 = P_1 + 2P_2$ and $Q'_1 = P_1 + P_2$. Then Q_1 and Q'_1 would no longer be redundant for the client, would be able to recover source contents P_1 and P_2, effectively doubling the throughput.

However, this is only half of the solution. Assuming the client targets high throughput and simultaneously sends *multiple parallel* Interests as in Fig. 2. Each of them would

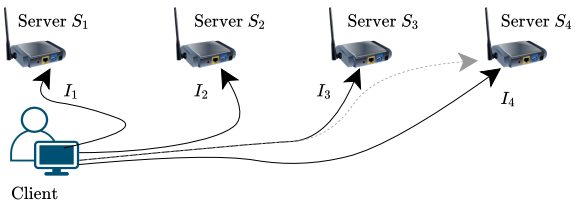

Fig. 2. What should the servers reply to each Interest sent in parallel?

This research was partly supported by Labex DigiCosme (project ANR11LABEX0045DIGICOSME) operated by ANR as part of the program "Investissement d'Avenir" Idex ParisSaclay (ANR11IDEX000302)

ideally bring back innovative content, but the question is how to guarantee this. To address this problem, in [5], [6], we proposed a construction called MILIC allowing one to build Multiple Interests for Linearly Independent Content. Motivated by the search for alternate solutions with potentially even better performance, we looked for generalizations of this construction. We could prove that a large class of generalizations are equivalent (*isomorphic*) to MILIC. The main contribution of this article is the proof for this fact in Theorem 4; this is of high interest because it limits the search space for alternate constructions, and the steps of the proof themselves shed light on properties the alternate solutions must have (or not have). Then, we indeed exhibit one different family of solutions.

The article is organized as follows: notations are introduced in Sect. II; a mathematical formulation of the problem is given in Sect. III; Sect. IV discusses its solutions, including MILIC; in Sect. V, a generic family of constructions (for potential solutions), is defined; Sect. VI and Sect. VII contains the proof that solutions from this family are included in MILIC; Sect. VIII shows alternative constructions; Sect. IX concludes.

II. Preliminary Notations and Definitions

Let $[\![k]\!] \stackrel{\text{def}}{=} \{1, 2, \dots k\}$ be the set of integers from 1 to k. F denotes a finite field with $|F| > 2$, $F^* \stackrel{\text{def}}{=} F \setminus \{0\}$, and F^n is the vector space of dimension $n \geqslant 1$ over F

$$F^n = \{\begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix} \mid x_1 \in F, x_2 \in F, \dots, x_n \in F\}.$$

The vector $e_i = \begin{pmatrix} \underbrace{0 \cdots 0}_{i-1 \text{ zeros}} & 1 & 0 \cdots 0 \end{pmatrix}$ is the i-th canonical vector of F^n.

Definition 1 (Encoding vector of a linear combination)**.** *If a coded packet Q is a linear combination of the source content $P_1, P_2, \dots P_n$, with $Q = \alpha_1 P_1 + \alpha_2 P_2 + \dots + \alpha_n P_n$, with $\forall i \in [\![n]\!], \alpha_i \in F$, then the encoding vector of Q is $v = \begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix}$.*

Definition 2 (Transversal)**.** *Let E be an arbitrary set. Consider a family of k subsets of E, denoted $\mathcal{A} = (\mathcal{A}_1, \dots, \mathcal{A}_k)$.*
A transversal *T of $\mathcal{A}$ is a set obtained by picking one element in each subset $\mathcal{A}_i$: $T = \{a_1, a_2, \dots a_k\}$ with $a_i \in \mathcal{A}_i, \forall i \in [\![k]\!]$.*

III. Multipath Retrieval and The Problem of Sets Ensuring Linearly Independent Transversals

Consider the content retrieval problem in the context of ICN illustrated in Fig. 2. A client wants to retrieve n segments $P_1, P_2, \dots, P_n$ of source content. The client sends n Interest packets $(I_i)_{i=1\dots n}$. The reply to the i-th interest I_i from any server will be a linear combination whose encoding vector comes from a predefined set $\mathcal{A}_i$. The problem considered in this paper is to build a good family of n sets $(\mathcal{A}_i)_{i \in [\![n]\!]}$. Such good family has to satisfy the following properties.

- Completeness (Fig. 2): When n coded packets with encoding vectors $v_1 \in \mathcal{A}_1, v_2 \in \mathcal{A}_2, \dots v_n \in \mathcal{A}_n$ are received by the client, it should be able to recover $P_1, P_2 \dots P_n$.
- High diversity (Fig. 1): If k coded packets with encoding vectors $v_1 \dots v_k$ are received, some potentially drawn from the same sets (*e.g.*, the following is possible: $v_1 \in \mathcal{A}_7, v_2 \in \mathcal{A}_7, v_3 \in \mathcal{A}_5, v_4 \in \mathcal{A}_5 \dots$), they should provide non-redundant information with high probability.

Our construction, MILIC (see Definition 3 in what follows), introduced in [5], satisfies both properties. In this work, because high diversity is arguably more difficult to formulate precisely for being inherently probabilistic, we ignore this aspect[1]. Instead, we focus on ensuring the completeness property. This can be recast as finding Sets Ensuring Linearly Independent Transversals (SELIT)[2] such as:

Problem 1 (SELIT problem)**.** *Given a finite field F, and the vector space F^n of dimension n over F:*
Find $\mathcal{A} = (\mathcal{A}_1, \dots, \mathcal{A}_n)$, a family of n disjoint subsets of F^n, such that any of their transversals (see definition 2) will always constitute a set of linearly independent vectors.

Clearly, a family $\mathcal{A}$ of sets solution of the SELIT problem satisfy the completeness property.

The SELIT problem is loosely related to the well known *index coding* problem [7]: given a server with a wireless broadcast link to several clients that already have some content, what is the minimum number of transmissions necessary to send the information, they want? And with which coding scheme? The index coding literature spawned results such as an equivalence with network coding [8], and capacity region results in the distributed case [9]. SELIT is simpler because it decouples the "capacity" aspect present in the index coding problem; we are unaware of existing constructions of the literature that would provide SELIT solutions[3].

IV. On MILIC and on Solutions of SELIT

One example of SELIT solution for $n = 2$ is in Ex. 1, where $\mathcal{A}_1 = \{\begin{pmatrix}1 & 2\end{pmatrix}, \begin{pmatrix}2 & 1\end{pmatrix}\}$, meaning that replies to Interest I_1 are $Q_1 = P_1 + 2P_2$ and $Q_1' = 2P_1 + P_2$; and any of these is linearly independent of any of the two possible replies for I_2 which are $Q_2 = P_1$ and $Q_2' = P_1 + P_2$.

Example 1.
$F = \mathrm{GF}(3), n = 2, k = 2, \mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ *with*
$\mathcal{A}_1 = \{\begin{pmatrix}1 & 2\end{pmatrix}, \begin{pmatrix}2 & 1\end{pmatrix}\}$
$\mathcal{A}_2 = \{\begin{pmatrix}1 & 0\end{pmatrix}, \begin{pmatrix}1 & 1\end{pmatrix}\}$
There are exactly 4 possible transversals:
- $T_1 = \{\begin{pmatrix}1 & 2\end{pmatrix}, \begin{pmatrix}1 & 0\end{pmatrix}\}$. $\det \left(\begin{smallmatrix}1 & 2\\ 1 & 0\end{smallmatrix}\right) = 1 \neq 0$
- $T_2 = \{\begin{pmatrix}1 & 2\end{pmatrix}, \begin{pmatrix}1 & 1\end{pmatrix}\}$. $\left|\begin{smallmatrix}1 & 2\\ 1 & 1\end{smallmatrix}\right| = 2 \neq 0$
- $T_1 = \{\begin{pmatrix}2 & 1\end{pmatrix}, \begin{pmatrix}1 & 0\end{pmatrix}\}$. $\left|\begin{smallmatrix}2 & 1\\ 1 & 0\end{smallmatrix}\right| = 1 \neq 0$
- $T_2 = \{\begin{pmatrix}2 & 1\end{pmatrix}, \begin{pmatrix}1 & 1\end{pmatrix}\}$. $\left|\begin{smallmatrix}2 & 1\\ 1 & 1\end{smallmatrix}\right| = 1 \neq 0$

Hence all 4 transversals are sets of 2 linearly independent vectors, and thus $\mathcal{A}$ is a SELIT solution.

[1] notice that high diversity is correlated with high cardinality of sets $\mathcal{A}_i$ in any case, so after finding large $\mathcal{A}_i$, one could check *a posteriori* their diversity

[2] The problem can also be generalized as $k-$SELIT for finding $k < n$ sets instead of n (with some uses for ICN): this is not explored in this article.

[3] but note that linear codes might map a SELIT to a k-SELIT solution.

Definition 3 (Multiple Interests for Linearly Independent Contents)**.** *MILIC is a construction with sets of encoding vectors* $(\mathcal{A}_i)_{i\in[\![n]\!]}$ *introduced in [5] as follows:* $\forall i \in [\![n]\!]$,

$$\mathcal{A}_i = \{(v_1, ..., v_n) \in F^n \mid v_i \neq 0 \text{ and } \forall j \in [\![i-1]\!], v_j = 0\}.$$

Example 2 (MILIC for $n = 3$)**.**

- $\mathcal{A}_1 = \{\begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} \mid a_1 \in F^*, a_2 \in F, a_3 \in F\}$
- $\mathcal{A}_2 = \{\begin{pmatrix} 0 & b_2 & b_3 \end{pmatrix} \mid b_2 \in F^*, b_3 \in F\}$
- $\mathcal{A}_3 = \{\begin{pmatrix} 0 & 0 & c_3 \end{pmatrix} \mid c_3 \in F^*\}$

Ex. 2 illustrates the MILIC construction [5] when $n = k = 3$. MILIC is one solution to the SELIT problem.

By construction, any transversal of the MILIC sets is linearly independent. Results in this article originate from attempts to generalize MILIC constructions with matrices where all coefficients are freely picked from predefined sets, with additional constraints, see Def. 6. They are for any field F with $|F| > 2$, that is, all excepted $GF(2)$. Theorem 4 states that a solution in this form is a MILIC solution up to a permutation of indices.

V. Investigated Families of Constructions

In this article, we investigate solutions in a form similar to MILIC: this section specifies what "similar" means, after introducing some additional definitions.

We first notice that if $\mathcal{A} = (\mathcal{A}_1, \ldots, \mathcal{A}_n)$ is a SELIT solution, one can take one arbitrary vector in each set $v_i \in \mathcal{A}_i$, and use the set of vectors $(v_i)_i$ as a new basis for coordinates. Accordingly, we introduce the concept of canonical family:

Definition 4 (Canonical family of sets)**.**
Let $\mathcal{A} = (\mathcal{A}_1, \ldots, \mathcal{A}_n)$ *be a family of* n *subsets of* F^n *:*
A canonical *family of sets is such that each canonical vector* e_i *is such that* $e_i \in \mathcal{A}_i$ *for all* $i \in [\![n]\!]$.

Definition 5 (Canonical and component-wise family of sets)**.** *A family* $\mathcal{A} = (\mathcal{A}_1, \ldots, \mathcal{A}_n)$ *of* n *subsets of* F^n *is a canonical and component-wise family of sets if subset* $\mathcal{A}_i, \in [\![n]\!]$ *is the union of* e_i *with the n-fold cartesian product of sets* $A_{i,j}$*, where* $A_{i,j} \subseteq F$ *are parameters of the construction. Such that* $\forall i \in [\![n]\!]$,

$$\mathcal{A}_i = \{ \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} \mid v_j \in A_{i,j} \ \forall j = 1 \ldots n\} \cup \{e_i\}. \quad (1)$$

A canonical and component-wise family of sets $\mathcal{A}$ is fully specified from the table $\mathcal{C}(\mathcal{A})$ of subsets $A_{i,j}$ denoted as

$$\mathcal{C}(\mathcal{A}) \stackrel{\text{def}}{=} \begin{Bmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,n} \end{Bmatrix}. \quad (2)$$

Example 3. *The MILIC construction* $\mathcal{A}^{\text{MILIC}}$ *in Def. 3 is a canonical and component-wise family, associated with the subsets* $A^{\text{MILIC}}_{i,j}$ *defined as follows:*

$$\begin{cases} A^{\text{MILIC}}_{i,i} = F^* & \text{if } i = j \\ A^{\text{MILIC}}_{i,j} = \{0\} & \text{if } i < j \\ A^{\text{MILIC}}_{i,j} = F & \text{if } i > j \end{cases}$$

and we have

$$\mathcal{C}(\mathcal{A}^{\text{MILIC}}) = \begin{Bmatrix} F^* & F & \cdots & F \\ \{0\} & F^* & \cdots & F \\ \vdots & \vdots & \ddots & F \\ \{0\} & \{0\} & \cdots & F^* \end{Bmatrix}$$

Notice that the Cartesian products defined in MILIC already contain the canonical unit vectors e_i.

Example 4. *Consider the MILIC construction* $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3)$ *with* $n = 3$ *with* $F = GF(3)$*; F can then be represented as* $F = \{0, 1, 2\}$*, and* $F^* = \{1, 2\}$*. Then*

$$\mathcal{C}(\mathcal{A}) = \begin{Bmatrix} \{1,2\} & \{0,1,2\} & \{0,1,2\} \\ \{0\} & \{1,2\} & \{0,1,2\} \\ \{0\} & \{0\} & \{1,2\} \end{Bmatrix}$$

Definition 6 (Canonical, Diagonal and Restricted Diagonal family of sets)**.** *Let* $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n)$ *be a canonical and component-wise family of* n *sets of* F*. We say that it is a* diagonal *family if there exists* $A_{i,j} \subset F$*,* $\forall i \in [\![n]\!]$*,* $\forall j \in [\![n]\!] \setminus \{i\}$ *such that*

$$\mathcal{C}(\mathcal{A}) = \begin{Bmatrix} F^* & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & F^* & \cdots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n,1} & A_{n,2} & \cdots & F^* \end{Bmatrix}_{n\times n}$$

It is a canonical and restricted diagonal *family if the off-diagonal sets* $(A_{i,j})$ *additionally satisfy:*

$$\forall i \in [\![n]\!], \forall j \in [\![n]\!] \setminus \{i\} : A_{i,j} \subset F^* \text{ or } A_{i,j} = \{0\}. \quad (3)$$

VI. Truncated Permutations

One important property of a canonical component-wise SELIT solution $\mathcal{A}$ for a given n is that one can extract smaller SELIT solutions $\mathcal{A}'$ for $n' < n$ by removing some sets and some coefficients of vectors of the sets. This is proven at the end of this section in Theorem 1 and will be used as the basis of the proof by induction of our main result in Sect. VII.

We first formally define what is meant by *extract*, through the definition of a *truncated permutation*.

A *truncated permutation* consists in applying some permutation ℓ_i to the elements of a vector, then truncating it to the first k elements. It is defined by $\ell_i \in [\![n]\!]$ for $\forall i \in [\![k]\!]$, such that $\forall j \in [\![k]\!] : \ell_i \neq \ell_j$ when $i \neq j$. The indices of the coefficients are specified by a tuple of k different indices $L = (\ell_1, \ell_2, \cdots, \ell_k)$. Later we use *mapping* for *truncated permutation*.

Definition 7 (Mapping of a vector)**.**
Let $v \in F^n$ be a vector $\begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix}$ and consider a tuple of k different indices $(\ell_1, \ell_2, \cdots, \ell_k)$. The mapping of v is

$$TP(v, \ell_1, \ell_2, \ldots, \ell_k) \stackrel{def}{=} \begin{pmatrix} v_{\ell_1} & v_{\ell_2} & \cdots & v_{\ell_k} \end{pmatrix} \in F^k \tag{4}$$

This definition can be extended to a family of subsets in a way that their diagonal structure is preserved.

Definition 8 (Mapping of a family of sets)**.**
Let $\mathcal{A} = (\mathcal{A}_1, \ldots, \mathcal{A}_n)$ be a family of sets of F^n.

The mapping *of $\mathcal{A}$, denoted $TP(\mathcal{A}, \ell_1, \ell_2, \ldots, \ell_k)$ with the indices $(\ell_1, \ell_2, \ldots, \ell_k)$ is the family of sets $\mathcal{A}'$ with:*

$$\begin{aligned} \mathcal{A}' &= (\mathcal{A}'_1, \ldots, \mathcal{A}'_k) \\ \mathcal{A}'_i &= \{TP(v, \ell_1, \ell_2, \ldots, \ell_k) \,|\, \forall v \in \mathcal{A}_{\ell_i}\}, \ \forall i \in [\![k]\!] \end{aligned} \tag{5}$$

When $\mathcal{A}$ is a component-wise family with coefficients from the sets $(A_{ij})_{i \in [\![n]\!], j \in [\![n]\!]}$, and we define a notation for the table of coefficients of the mapping:

$$\mathcal{P}(\mathcal{A}, \ell_1, \ell_2, \ldots, \ell_k) \stackrel{def}{=} \mathcal{C}(TP(\mathcal{A}, \ell_1, \ell_2, \ldots, \ell_k))$$

$$\mathcal{P}(\mathcal{A}, \ell_1, \ell_2, \ldots, \ell_k) = \begin{Bmatrix} A_{\ell_1\ell_1} & A_{\ell_1\ell_2} & \cdots & A_{\ell_1\ell_k} \\ A_{\ell_2\ell_1} & A_{\ell_2\ell_2} & \cdots & A_{\ell_2\ell_k} \\ \vdots & \vdots & \ddots & \vdots \\ A_{\ell_k\ell_1} & A_{\ell_n\ell_2} & \cdots & A_{\ell_k\ell_k} \end{Bmatrix} \tag{6}$$

Example 5. *For MILIC with $n = 3$ (defined in example 4), we have the following examples of mapping:*

$$\mathcal{P}(\mathcal{A}, 1, 2, 3) = \begin{Bmatrix} F^* & F & F \\ \{0\} & F^* & F \\ \{0\} & \{0\} & F^* \end{Bmatrix} \text{ with } \begin{matrix} F = \{0, 1, 2\} \\ F^* = \{1, 2\} \end{matrix}$$

$$\mathcal{P}(\mathcal{A}, 1, 2) = \begin{Bmatrix} F^* & F \\ \{0\} & F^* \end{Bmatrix} \text{ and } \mathcal{P}(\mathcal{A}, 2, 1) = \begin{Bmatrix} F^* & \{0\} \\ F & F^* \end{Bmatrix}.$$

Example 6. *When $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n)$ is a MILIC construction then for any set of k indices $\ell_1, \ell_2, \ldots, \ell_k$ with $\ell_1 < \ell_2 \ldots < \ell_k$, we have:*

$$\mathcal{P}(\mathcal{A}, \ell_1, \ell_2, \ldots, \ell_k) = \begin{Bmatrix} F^* & F & F & \cdots & F \\ \{0\} & F^* & F & \cdots & F \\ \{0\} & \{0\} & F^* & \cdots & F \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \{0\} & \{0\} & \{0\} & \cdots & F^* \end{Bmatrix} \tag{7}$$

Theorem 1 (Mapping Theorem for Canonical Solutions)**.** *Let $\mathcal{A}$ be a canonical family of sets, that is a SELIT solution to the SELIT problem for n, and $\ell_1, \ell_2, \ldots, \ell_k$ be a sequence of distinct k indices in $[\![n]\!]$.*

Then $\mathcal{A}' = TP(\mathcal{A}, \ell_1, \ell_2, \ldots, \ell_k)$ is a family of k sets of F^k and a SELIT solution to the SELIT problem for dimension k.

Proof. By contradiction: assume $\mathcal{A}'$ is not a SELIT solution. It implies that there exist $w_i \in \mathcal{A}'_i, \forall i \in [\![k]\!]$ and $\alpha \in F^k$, $\alpha \neq 0$ such that $\alpha_1 w_1 + \alpha_2 w_2 + \ldots + \alpha_k w_k = 0$. $\mathcal{A}'$ is a family of subsets that are mapping of $\mathcal{A}$, which implies that each vector $w_i \in \mathcal{A}'_i$ is a mapping of some vector $v_i \in \mathcal{A}_i$ from (7). Then

$$\sum_{i \in [\![k]\!]} \alpha_i v_i = \begin{pmatrix} \underbrace{0 & 0 & \cdots & 0}_{k \text{ zeros}} & x_{k+1} & x_{k+2} & \cdots & x_n \end{pmatrix}$$

up to a permutation of the entries. Since $\mathcal{A}$ is a canonical family of sets, each $\mathcal{A}_j$ includes the canonical vector e_j, and then

$$\sum_{\forall i \in [\![k]\!]} \alpha_i v_i - \sum_{\forall i \in \{k+1, \ldots, n\}} x_i e_i = 0$$

which is a contradiction with the fact that $\mathcal{A}$ is a SELIT solution. Hence $\mathcal{A}'$ must be a SELIT solution. □

VII. Properties of Canonical and Restricted Diagonal Solutions of the SELIT problem

In this entire section, we consider one solution of the SELIT problem $\mathcal{A}$ that is canonical, component-wise, and restricted diagonal. Recall that by the previous definitions we can write:

$$\begin{cases} \mathcal{A} = (\mathcal{A}_1, \ldots, \mathcal{A}_n), \text{ with } \forall i \in [\![n]\!] : e_i \in \mathcal{A}_i \\ \mathcal{C}(\mathcal{A}) = \begin{Bmatrix} F^* & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & F^* & \ddots & \vdots \\ \vdots & \ddots & \ddots & A_{n-1,n} \\ A_{n,1} & \cdots & A_{n,n-1} & F^* \end{Bmatrix} \\ \text{with } \forall i \in [\![n]\!], \forall j \in [\![n]\!] \setminus \{i\} : A_{i,j} \subset F^* \text{ or } A_{i,j} = \{0\} \end{cases} \tag{8}$$

and we will use these definitions.

Let ℓ_1, ℓ_2 be two different integers in $[\![n]\!]$. The 2×2 mapping, $\mathcal{P}(\mathcal{A}, \ell_1, \ell_2) = \begin{Bmatrix} F^* & A_{\ell_1,\ell_2} \\ A_{\ell_2,\ell_1} & F^* \end{Bmatrix}$ involves only A_{ℓ_1,ℓ_2} and A_{ℓ_2,ℓ_1}, that have an important property:

Lemma 1 (Non-Zero Set Exclusion)**.** *Consider $n > 1$. Let $\mathcal{A}$ be a SELIT solution as in* (8)*. Let ℓ_1, ℓ_2 be two different integers in $[\![n]\!]$.*

Then $A_{\ell_1,\ell_2} = \{0\}$ or $A_{\ell_2,\ell_1} = \{0\}$ or both are $\{0\}$

Proof. By contradiction: assume that both $A_{\ell_1,\ell_2} \neq \{0\}$ and $A_{\ell_2,\ell_1} \neq \{0\}$. Then from (8), $A_{\ell_1,\ell_2} \subset F^*$, and $A_{\ell_2,\ell_1} \subset F^*$, hence $0 \notin A_{\ell_1,\ell_2}$ and $0 \notin A_{\ell_2,\ell_1}$. Let $\mathcal{A}' = (\mathcal{A}'_1, \ldots, \mathcal{A}'_2)$ be the mapping of $\mathcal{A}$, *i.e.* $\mathcal{A}' = \mathcal{P}(\mathcal{A}, \ell_1, \ell_2)$. We know that:

$$\mathcal{C}(\mathcal{A}') = \mathcal{P}(\mathcal{A}, \ell_1, \ell_2) = \begin{Bmatrix} F^* & A_{\ell_1,\ell_2} \\ A_{\ell_2,\ell_1} & F^* \end{Bmatrix}$$

Take arbitrary values $a_1 \in \mathcal{A}_{\ell_1,\ell_2}$, and $a_2 \in \mathcal{A}_{\ell_2,\ell_1}$, necessarily non-zero, and let $v = \begin{pmatrix} a_2 & a_1 \end{pmatrix}$. Then $v \in \mathcal{A}'_1$ and $v \in \mathcal{A}'_2$, and v is linearly dependent with itself.

Thus $\mathcal{A}'$ is not a SELIT solution and by theorem 1, neither is $\mathcal{A}$. This is a contradiction, therefore the initial assumption that both A_{ℓ_1,ℓ_2} and A_{ℓ_2,ℓ_1} are not $\{0\}$ must be false, hence the lemma. □

This lemma allows us to further characterize SELIT solutions in binary relations.

Definition 9 (Binary Relations from $\mathcal{A}$)**.** *Consider $\mathcal{A}$ as in* (8)*, $\ell_1 \in [\![n]\!]$ and $\ell_2 \in [\![n]\!]$ with $\ell_1 \neq \ell_2$. We can introduce the binary relations on ℓ_1 and ℓ_2, depending on whether A_{ℓ_1,ℓ_2} and A_{ℓ_2,ℓ_1} are $\{0\}$ (we also indicate how $P \overset{def}{=} \mathcal{P}(\mathcal{A}, \ell_1, \ell_2)$ looks like):*

$$\ell_1 \overset{\mathcal{A}}{\prec} \ell_2 \text{ if only } A_{\ell_2,\ell_1} = \{0\};\ \textit{thus: } P = \begin{Bmatrix} F^* & A_{\ell_1,\ell_2} \\ \{0\} & F^* \end{Bmatrix} \tag{9}$$

$$\ell_2 \overset{\mathcal{A}}{\prec} \ell_1 \text{ if only } A_{\ell_1,\ell_2} = \{0\};\ \textit{thus: } P = \begin{Bmatrix} F^* & \{0\} \\ A_{\ell_2,\ell_1} & F^* \end{Bmatrix} \tag{10}$$

$$\ell_2 \overset{\mathcal{A}}{\sim} \ell_1 \text{ else when both are } \{0\};\ \textit{thus: } P = \begin{Bmatrix} F^* & \{0\} \\ \{0\} & F^* \end{Bmatrix} \tag{11}$$

$$\text{Observe that: } \ell_1 \overset{\mathcal{A}}{\not\prec} \ell_2 \iff A_{\ell_1,\ell_2} = \{0\} \tag{12}$$

The overscript $\mathcal{A}$ makes clear that the binary relations depend on $\mathcal{A}$. To simplify, when there is no ambiguity from the context, we will write $\prec$ instead of $\overset{\mathcal{A}}{\prec}$.

We now start with a lemma on the determinant of matrices.

Lemma 2. *Consider the sets $B_{i,j}$, $j \in [\![k]\!], i \in [\![j-1]\!]$, the sets $D_i \subset F^*$, $i \in [\![k]\!] \subset F$, and the set set of matrices $\mathcal{M}_k$ defined as*

$$\mathcal{M}_k = \left\{ \begin{pmatrix} d_1 & b_{1,2} & b_{1,3} & \cdots & b_{1,k-1} & b_{1,k} \\ c_1 & d_2 & b_{2,3} & \cdots & b_{2,k-1} & b_{2,k} \\ 0 & c_2 & d_3 & \cdots & b_{3,k-1} & b_{3,k} \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & d_{k-1} & b_{k-1,k} \\ 0 & 0 & 0 & \cdots & c_{k-1} & d_k \end{pmatrix} \right.$$
$$\left. \mid b_{i,j} \in B_{i,j}, \forall j \in [\![k]\!], \forall i \in [\![j-1]\!], \right.$$
$$\left. \text{and } d_i \in D_i, \forall i \in [\![k]\!], \text{ and } c_i \in F^*, \forall i \in [\![k-1]\!] \quad \right\},$$

Then there exists a matrix $M \in \mathcal{M}_k$ such that $\det(M) \neq 0$

Proof. For $k = 2$: consider a 2×2 matrix $M_2 \in \mathcal{M}$

$$M_2 = \begin{pmatrix} d_1 & b_{1,2} \\ c_1 & d_2 \end{pmatrix}$$

where d_1, d_2, arbitrarily chosen from the predefined sets $D_i \subset F^*$, $i = 1, 2$, are necessarily non-zero; $b_{1,2} \in B_{1,2}$ is also arbitrary chosen, and might be zero. If $b_{1,2} = 0$ then $\det(M_2) \neq 0$. Otherwise if $b_{1,2} \neq 0$ then for each value of c_1 the $\det(M_2)$ will have a different value, thus there exists c_1 with $\det(M_2) \neq 0$. Thus the result for $k = 2$.

We prove the general case by induction. Assume the property is proven for $\mathcal{M}_{k-1}$ for some $k > 1$. The determinant of all $k \times k$ matrices $M_k \in \mathcal{M}_k$ can be written as

$$\det(M_k) = d_1 \det(M'_{k-1}) - c_1 \det(Q)$$

where $M'_{k-1} \in \mathcal{M}_{k-1}$ and Q is some $(k-1) \times (k-1)$ matrix. Neither M'_{k-1} nor Q involves the coefficient c_1. From the induction assumption, we know that there exists $M'_{k-1} \in \mathcal{M}_{k-1}$ such that $\det(M'_{k-1}) \neq 0$. The matrix M_k is built from M'_{k-1}, selecting arbitrarily the remaining coefficients from their possible sets, except c_1. If $\det(Q) = 0$, then $\forall c_1 \in F^*$, $\det(M_k) \neq 0$. If $\det(Q) \neq 0$, $\det(M_k)$ will take as many different values as c_1 and $\exists\ c_1 \in F^*$ such that $\det(M_k) \neq 0$. Assuming the property the lemma true for $\mathcal{M}_{k-1}$, it is thus also true for $\mathcal{M}_k$. Since it is true for $\mathcal{M}_2$, the lemma is proven. □

Theorem 2. *Let $\mathcal{A}$ be a SELIT solution as in* (8)*. Let $k > 1$. Consider a sequence of k indices $L = \{\ell_1, \ell_2, \ldots, \ell_k\}$, such that $\ell_1 \prec \ell_2, \ell_2 \prec \ell_3, \ldots,$ and $\ell_{k-1} \prec \ell_k$, then $\ell_k \not\prec \ell_1$.*

Proof. By the property of $\prec$ in (9) and (10), $\ell_1 \prec \ell_2$ and $\ell_2 \prec \ell_1$ exclude each other, and the theorem is proven for $k = 2$. We will prove the theorem for $k > 2$, by induction, assuming that it was proven for $2, 3, \ldots, k-1$.

Consider a sequence of k indices $\mathrm{L} = \{\ell_1, \ell_2, \ldots, \ell_k\}$, that satisfies $\ell_1 \prec \ell_2, \ell_2 \prec \ell_3, \ldots,$ and $\ell_{k-1} \prec \ell_k$.
For any $i \in [\![k]\!]$ and $j \in [\![k]\!]$ with $i < j$ and $(i,j) \neq (1,k)$, we can apply the induction hypothesis for $\ell_i \prec \ell_{i+1}, \ell_{i+1} \prec \ell_{i+2}, \ldots,$ and $\ell_{j-1} \prec \ell_j$, which forms a chain of $k' = j - i + 1 \leq k - 1$ binary relations. Thus $\ell_j \not\prec \ell_i$ and from (12),

$$\forall i \in [\![k]\!], \forall j \in [\![k]\!] \text{ with } i < j, (i,j) \neq (1,k) : \mathcal{A}_{\ell_j,\ell_i} = \{0\}$$

Notice that these are all the elements in the lower triangle of the table $\mathcal{P}(\mathcal{A}, \ell_1, \ell_2, \ldots, \ell_k)$, and are $\{0\}$, except for the diagonal and for A_{ℓ_k,ℓ_1}.

$$P_{\mathrm{L}} = \left\{ \begin{matrix} F^* & A_{\ell_1,\ell_2} & \cdots & A_{\ell_1,\ell_{k-1}} & A_{\ell_1,\ell_k} \\ \{0\} & F^* & \cdots & A_{\ell_2,\ell_{k-1}} & A_{\ell_2,\ell_k} \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ \{0\} & \{0\} & \ddots & F^* & A_{\ell_{k-1},\ell_k} \\ A_{\ell_k,\ell_1} & \{0\} & \cdots & \{0\} & F^* \end{matrix} \right\}. \tag{13}$$

Consider $\mathcal{A}' = \mathrm{TP}(\mathcal{A}, \ell_1, \ell_2, \ldots, \ell_k)$ a mapping of the SELIT solution $\mathcal{A}$. Consider the set of matrices $\mathcal{Q}_{\mathcal{A}'}$ constructed from vectors obtained by picking one vector in each set of $\mathcal{A}'$. Each matrix $M = (m_{i,j})$ in $\mathcal{Q}_{\mathcal{A}'}$ is obtained by picking a coefficient in the table P_{L} in (13) at the corresponding position. Its determinant can be written as:

$$\det(M) = m_{1,1} \det(G) - m_{k,1} \det(H)$$

where $m_{1,1} \in F^*$, G is a triangular matrix (with diagonal elements in F^*), $m_{k,1} \in A_{\ell_k,\ell_1}$ and H is a matrix in form of $\mathcal{M}_{k-1}$ from Lemma 2. This implies that coefficients present in matrix H can be selected such that $\det(H) \neq 0$. Moreover, $\det(G) \neq 0$ as $\det(G)$ is the product of non-zero elements. If $m_{k,1}$ is not zero, then there exists a value $m_{1,1} = m_{k,1}\det(H)/\det(G) \in F^*$ such that $\det(M) = 0$, and then the corresponding vectors of M would be linearly dependent, which contradicts the fact that $\mathcal{A}'$ (thus $\mathcal{A}$) is a SELIT solution. Thus $m_{k,1}$ must always be 0, hence, $0 \in A_{\ell_k,\ell_1}$, and from (8), this implies that $A_{\ell_k,\ell_1} = \{0\}$.

From (12), this now implies that $\ell_k \not\prec \ell_1$ and this concludes the proof by induction, hence the theorem. □

Theorem 3. *Any canonical and restricted diagonal family $\mathcal{A} = (\mathcal{A}_1, \ldots, \mathcal{A}_n)$ that is a solution of SELIT, is included in a family $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_n)$ which is a permutation of lines and rows of the MILIC construction for n (where "included" means that $\forall i \in [\![n]\!], \mathcal{A}_i \subset \mathcal{B}_i$)*

Proof. The binary relation $\prec$ is acyclic as a direct consequence of Theorem 2. We can use classical results to build a total order embedding it, see for instance [10]: let $\lhd$ be the transitive closure of $\prec$; the transitive closure of an acyclic relation is irreflexive (see [10]). Now, by the Szpilrajn extension theorem [11], for a transitive and irreflexive relation, there exists a total order, that includes it, denoted $\lll$. We can reorder the indices $[\![n]\!]$ as a sequence $\mathrm{L} = \{\ell_1, \ell_2, \ldots, \ell_n\}$ using the total order such that $\ell_1 \lll \ell_2 \lll \ldots \ell_{n-1} \lll \ell_n$. The mapping of $\mathcal{A}$ with L (actually a permutation) is given by:

$$\mathcal{P}(\mathcal{A}, \ell_1, \ell_2, \ldots, \ell_n) = \begin{Bmatrix} F^* & A_{\ell_1,\ell_2} & \cdots & A_{\ell_1,\ell_n} \\ A_{\ell_2,\ell_1} & F^* & \cdots & A_{\ell_2,\ell_n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{\ell_n,\ell_1} & A_{\ell_n,\ell_2} & \cdots & F^* \end{Bmatrix}$$

Consider any $\ell_i \in [\![n]\!], \ell_j \in [\![n]\!]$ and $\ell_i \neq \ell_j$. Assume that $A_{\ell_i,\ell_j} \neq \{0\}$. From (12), $A_{\ell_i,\ell_j} \neq \{0\}$ iff $\ell_i \prec \ell_j$. Then $\ell_i \prec \ell_j$ implies that $\ell_i \lhd \ell_j$ and consequently $\ell_i \lll \ell_j$. And then $i < j$ (because $\ell_i \lll \ell_j \iff i < j$), and this element A_{ℓ_i,ℓ_j} must be in the upper triangle.

As a consequence, if ℓ_i, ℓ_j, correspond to the lower triangle (*e.g.* $j < i$) then $A_{\ell_i,\ell_j} = \{0\}$. This proves that the family $\mathcal{A}' = \mathcal{P}(\mathcal{A}, \mathrm{L})$ has a triangular matrix, hence is included in the MILIC construction for n, shown in (7). $\mathcal{A}'$ is obtained through a permutation of $\mathcal{A}$, hence the theorem, with $\mathcal{B}$ obtained through the inverse permutation of the MILIC construction for n. □

Now the most general form of the theorem is obtained by no longer considering "restricted diagonal" families, but any "diagonal" families:

Theorem 4. *Any canonical and diagonal family $\mathcal{A} = (\mathcal{A}_1, \ldots, \mathcal{A}_n)$ that is a solution of SELIT, is included in a family $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_n)$ which is a permutation of lines and rows of the MILIC construction for n (where "included" means that $\forall i \in [\![n]\!], \mathcal{A}_i \subset \mathcal{B}_i$).*

Proof. Let $(A_{i,j})$ be the sets of coefficients in $\mathcal{C}(A)$. Let $\mathcal{A}'$ be the family defined from its sets of coefficients $(A'_{i,j})$ from $\mathcal{C}(A')$ selected as follows:
$\forall i \in [\![n]\!], \forall j \in [\![n]\!]$: if $A_{i,j} \neq \{0\}$ then $A'_{i,j} \stackrel{\text{def}}{=} A_{i,j} \setminus \{0\}$ otherwise $A'_{i,j} \stackrel{\text{def}}{=} \{0\}$.

$\mathcal{A}'$ is now a restricted diagonal solution, hence Theorem 3 can be applied, and $\mathcal{B}$, a permutation of a MILIC construction can be found such that $\mathcal{A}'$ is included in $\mathcal{B}$. Now the only difference between $\mathcal{A}'$ and $\mathcal{A}$, is that vectors of transversals of $\mathcal{A}$ may have 0 in coefficient positions where vectors of transversals of $\mathcal{A}'$ might not: but then they would still be included in $\mathcal{B}$, hence $\mathcal{A}$ is included in $\mathcal{B}$, hence the theorem. □

VIII. Alternate Algebraic SELIT solutions

The details of the proofs in this article gave us insights on the reasons why our constructions have specific structures. Consider $M = \begin{pmatrix} x & a \\ b & c \end{pmatrix}$. We can write $\det(M) = 0 \iff x = abc^{-1}$. When x can be selected freely from F^*, we can make $\det(M) = 0$, unless a, b or c must be zero. Notice alternately that if a, b, and/or c can take several values, the generalizations of the Cauchy-Davenport theorem show that abc^{-1} will take even more values, hence making $\det(M) = 0$ becomes easier, unless, for instance, they are in a subgroup.

This is the insight that leads us to alternate families of SELIT solutions:

Theorem 5. *Let $H_0 = \{0\}$, and let $H_1 \subset \ldots \subset H_n$ be different subfields of F and denote $\forall i \in [\![n]\!], C_i \stackrel{def}{=} H_i \setminus H_{i-1}$. Let $\mathcal{A}$ be the "Matryoshka" set family[a] such that $\mathcal{C}(\mathcal{A})$ is:*

$$\begin{pmatrix} C_n & C_{n-1} & C_{n-1} & \cdots & C_{n-1} & C_{n-1} & C_{n-1} & C_{n-1} \\ C_{n-1} & C_{n-1} & C_{n-2} & \cdots & C_{n-2} & C_{n-2} & C_{n-2} & C_{n-2} \\ C_{n-1} & C_{n-2} & C_{n-2} & \cdots & C_{n-3} & C_{n-3} & C_{n-3} & C_{n-3} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \\ C_{n-1} & C_{n-2} & C_{n-3} & \cdots & C_4 & C_3 & C_3 & C_3 \\ C_{n-1} & C_{n-2} & C_{n-3} & \cdots & C_3 & C_3 & C_2 & C_2 \\ C_{n-1} & C_{n-2} & C_{n-3} & \cdots & C_3 & C_2 & C_2 & C_1 \\ C_{n-1} & C_{n-2} & C_{n-3} & \cdots & C_3 & C_2 & C_1 & C_1 \end{pmatrix}$$

Then $\mathcal{A}$ is a SELIT solution.

[a] Note that given a finite field F, the "Matryoshka" construction is only possible until a fixed n, which depends on existing F subfields. But for any n, one can select a finite field F that allows the construction,.

Proof. Pick an arbitrary transversal from $\mathcal{A}$, and write the matrix M_n of its vectors. Then one can write:
$\det(M_n) = a_n \det(M_{n-1}) + \ldots + (-1)^{n-1} a_1 \det(B_1)$ where

$a_n \in c_n$, $a_{n-1} \in c_{n-1} \dots a_1 \in c_{n-1}$ and M_{n-1} is a matrix with a similar form as M_n of size $(n-1) \times (n-1)$ and $B_1, \dots B_{n-1}$ are matrices with coefficients in H_{n-1}.

$\det(M_n) = a_n \det(M_{n-1}) + h$ where $h \in H_{n-1}$.

If $\det(M_{n-1}) \in H_{n-1}$ and $\det(M_{n-1}) \neq 0$, then it is not possible to take $a_n = -\det(M_{n-1})^{-1} h \in H_{n-1}$ because $a_{n+1} \in c_n = H_n \setminus H_{n-1}$, hence $\det(M_n) \neq 0$, and still $\det(M_n) \in H_n$. This is the basis for a proof of induction that $\det(M_n) \neq 0$, knowing that for $n = 1$ the property is true. And hence $\mathcal{A}$ is a SELIT solution. □

IX. Conclusion

In this article, we have introduced a new formal problem: the SELIT problem. We have proven that a large class of solutions for $|F| \geq 3$, where one can pick the coefficient vectors from fixed sets and with additional constraints, are essentially a version of the MILIC construction previously introduced. One example of an alternate construction was provided. Open questions and future work include the following: what can be said for families of solutions that are not canonical (which do not include the canonical vectors)? What about k-SELIT where the family $\mathcal{A}$ contains only $k < n$ elements? Can other constructions be proposed (using coding theory results)?

References

[1] V. Jacobson, D. K. Smetters, J. D. Thornton, M. F. Plass, N. H. Briggs, and R. L. Braynard. Networking named content. In *Proc. ACM CONEXT*, pages 1–12. ACM, 2009.

[2] Rudolf Ahlswede, Ning Cai, Shuo-yen Robert Li, and Raymond W. Yeung. Network Information Flow. *IEEE Transactions on Information Theory*, 46(4):1204–1216, 2000.

[3] M.-J. Montpetit, C. Westphal, and D. Trossen. Network coding meets information-centric networking: An architectural case for information dispersion through native network coding. *MobiHoc*, pages 31–36, 2012.

[4] Kazuhisa Matsuzono, Hitoshi Asaeda, and Cedric Westphal. Network Coding for Content-Centric Networking / Named Data Networking: Requirements and Challenges. Internet-Draft draft-irtf-nwcrg-nwc-ccn-reqs-04, IETF, Sep 2020. Work in Progress.

[5] H Malik, C Adjih, C Weidmann, and M Kieffer. MICN: a Network Coding Protocol for ICN with Multiple Distinct Interests per Generation. *arXiv preprint arXiv:2007.01128*, 2020.

[6] H. Malik, C. Adjih, M. Kieffer, and C. Weidmann. Analysis of the properties of NetcodICN protocols. In *CORES 2020*, Lyon, France, September 2020.

[7] Y. Birk and T. Kol. Informed-source coding-on-demand (iscod) over broadcast channels. In *IEEE INFOCOM*, pages 1257–1264, 1998.

[8] M. Effros, S. El Rouayheb, and M. Langberg. An equivalence between network coding and index coding. *IEEE Trans. Inf. Theory*, 61(5):2478–2487, 2015.

[9] Y. Liu, P. Sadeghi, F. Arbabjolfaei, and Y. Kim. Capacity theorems for distributed index coding. *IEEE Transactions on Information Theory*, 66(8):4653–4680, 2020.

[10] Theodore C Bergstrom. Maximal elements of acyclic relations on compact sets. *Journal of Economic Theory*, 10(3):403 – 404, 1975.

[11] Edward Szpilrajn. Sur l'extension de l'ordre partiel. *Fundamenta mathematicae*, 1(16):386–389, 1930.