Tight Information Theoretic Converse Results for some Pliable Index Coding Problems

Tang Liu and Daniela Tuninetti
University of Illinois at Chicago, Chicago, IL 60607 USA, Email: tliu44, danielat@uic.edu

***Abstract*—This paper studies the Pliable Index CODing problem (PICOD), which models content-type distribution networks. In the PICOD(t) problem there are m messages, n users and each user has a distinct message side information set, as in the classical Index Coding problem (IC). Differently from IC, where each user has a pre-specified set of messages to decode, in the PICOD(t) a user is "pliable" and is satisfied if it can decode any t messages that are not in its side information set. The goal is to find a code with the shortest length that satisfies all the users. This flexibility in determining the desired message sets makes the PICOD(t) behave quite differently compared to the IC, and its analysis even more challenging.**

This paper mainly focuses on the *complete–*S PICOD(t) with m messages, where the set $S \subset [m]$ contains the sizes of the side information sets, and the number of users is $n = \sum_{s\in S} \binom{m}{s}$, with no two users having the same side information set. Capacity results are shown for: (i) the *consecutive* complete–S PICOD(t), where $S = [s_{\min} : s_{\max}]$ for some $0 \leqslant s_{\min} \leqslant s_{\max} \leqslant m-t$, and (ii) the *complement-consecutive* complete–S PICOD(t), where $S = [0 : m-t] \backslash [s_{\min} : s_{\max}]$, for some $0 < s_{\min} \leqslant s_{\max} < m-t$. The novel converse proof is inspired by combinatorial design techniques and the key insight is to consider all messages that a user can eventually decode successfully, even those in excess of the t required ones. This allows one to circumvent the need to consider all possible desired message set assignments at the users in order to find the one that leads to the shortest code length. The core of the novel proof is to solve the *critical* complete–S PICOD(t) with $m = 2s + t$ messages and $S = \{s\}$, by showing the existence of a user who can decode $s+t$ messages regardless of the desired message set assignment. All other tight converse results for the complete–S PICOD(t) can be deduced from this critical case. The converse results show the information theoretic optimality of simple linear coding schemes. By similar reasoning, all complete–S PICOD(t) where the number of messages is $m \leqslant 5$ can be fully characterized. In addition, tight converse results are also shown for the PICOD(1) with circular-arc network topology hypergraph.

I. Introduction

A. Motivation

The broadcast channel with message side information at the receivers has became a critical model to understand the full potential of wireless communication networks as it models, for example, the downlink of the two-way relay channel [2]. Not even the capacity of the general broadcast channel without receiver side information is known. Therefore, some practically motivated and reasonably simple models are of interest when message side information at the receivers is considered. Index coding (IC) is one such model. First proposed in [4] when considering satellite communication, the IC consists of one transmitter with m independent messages to be delivered to n users through an error-free broadcast link. Each user has some messages as side information available to it and needs to reliably decode some messages that are not in its side information set; the desired messages for each user are pre-determined. In IC, one asks what is the minimum number of transmissions (i.e., minimum code length) such that every user is able to decode its desired messages successfully [3]. Compared to the general broadcast channel with side information at the users, the IC appears simple because: 1) the channel is noiseless, and 2) the side information sets are proper subsets of the whole message set. The IC focuses on the benefits / transmitter encoding opportunities brought by the different side information sets at the users. However, the general IC is still open. When one restricts attention to linear codes, the optimal code length is fully characterized by the so-called *minrank* problem, which is NP-complete in general [3]. Interestingly, it has been shown that every unicast network coding instance has an "equivalent" IC instance, meaning that, the network coding instance is solvable if and only if its corresponding IC instance is solvable [10], [1, Theorem 10.1]. This proves that the IC, which is a special network coding problem, is in fact equivalent to the general unicast network coding problem. This implies that for the IC, as for network coding, linear schemes are not sufficient [9] and non-Shannon type of inequality are necessary [13].

The IC problem models scenarios where the transmitter can do encoding based on the side information sets and on fixed desired message sets for the users. In practice, there may be flexibility in choosing the desired message sets. For example, in a music streaming service, users do not know which song will be played next; they are usually only guaranteed that it will be one from a certain group and not repeated. In online advertisement systems, the clients do not require a specific advertisement to see; it is the distributor who chooses what will be put on the clients' screens; the distributor might want to avoid repeating the same advertisement at the same client, as it might decrease the client's satisfaction. These scenarios can be modeled as a variant of the IC where the users are satisfied by *any* message that is not in its side information set, instead of a specific one as in the original IC setting. The transmitter thus has the freedom to choose the messages conveyed to the users so to minimize the transmission duration, or code length.

In this paper, we study this variant of IC known as Pliable Index CODing (PICOD), firstly proposed in [5]. The

This work was presented in part at ITW 2017 and ITW 2018. The work of the authors was partially funded by NSF under award number 1527059. The contents of this article are solely the responsibility of the author and do not necessarily represent the official views of the NSF.

PICOD(t) and the IC share many attributes. In the PICOD(t), one still has a single transmitter with m message and n users with message side information. The transmitter and users are connected via a shared noiseless rate-limited broadcast channel. The only major difference is that for the PICOD(t) the desired message sets at the users are not pre-determined and each user is satisfied whenever it can decode *any* t messages not in its side information set. This provides the transmitter more encoding opportunities, as it now encodes based on its own choice of desired message sets for the users, by knowing the message side information sets at each user. The goal in the PICOD(t) is to find the desired message set assignment that leads to the smallest possible code length.

B. Past Work on PICOD

As one would expect, the extra freedom of choosing the desired message sets in the PICOD(t) significantly reduces the number of transmissions / code length compared to the classical IC with the same number of messages, number of users, and message side information sets. In [5], when all side information sets are of size $s \leqslant m-t$, it was shown that there exits a code of length $O(\min\{t\log(n), t+\log^2(n)\})$ for the PICOD(t). When there is no constraint on the size of side information sets, and $m = O(n^\delta)$ for some constant positive δ, a code length of $O(\min\{t\log^2(n), t\log(n)+\log^3(n)\})$ is achievable [5]. Recently in [12], a deterministic polynomial time algorithm was proposed to achieve a code length of $O(\log^2(n))$ for $t=1$ and of $O(t\log(n)+\log^2(n))$ otherwise.

An interesting model proposed in [5] is the so-called *oblivious* PICOD(t). In the oblivious PICOD(t) the transmitter does not know the specific side information sets at the users. The transmitter only has knowledge of the sizes of the side information sets. In [5], [7] the authors proved that for the oblivious PICOD(t) at least a fraction $1/e$ of the remained unsatisfied users can be satisfied at each new transmission. This shows that there exists an achievable scheme where the code length is the logarithm of the number of users in the system, which is an exponential improvement in the number of transmissions compared to the IC.

Known achievable schemes for the PICOD(t) are based on linear codes only, and very few converse results are available. To the best of our knowledge, all converse proofs show bounds under the constraint that the code used is linear. For the oblivious PICOD(t), the optimal code length under the restriction that the transmitter can only use linear schemes is shown in [5, Theorem 9]. In [12], the authors provide a worst case instance that needs $\Omega(\log(n))$ code length for linear codes.

C. Contributions

In this paper we derive tight information theoretic converse bounds for some PICOD(t) problems based on the structure of the side information sets, namely: (i) the *complete-*S PICOD(t), and (ii) the PICOD(t) with a *circular-arc network topology hypergraph*.

The complete–S PICOD(t), where S is a subset of $[0 : m-t]$ (where m is the number of messages at the transmitter and t the number of messages to be decoded), is a system where all side information sets / users with size indexed by S are present. We say that S is *consecutive* if $S = [s_{\min} : s_{\max}]$ for some $0 \leqslant s_{\min} \leqslant s_{\max} \leqslant m-t$, which is also known as the oblivious PICOD(t) in [5]. In [5] the authors derived tight converse bounds for the oblivious PICOD(t) when the coding scheme is restricted to be a linear code. *In this work, we aim to provide tight information theoretic converse bounds, i.e., without any restriction on the coding scheme being used, on the same model. Our complete–*S *PICOD*(t) *setting actually includes and expands on the oblivious PICOD*(t) *setting studied in [5], and our results show the unrestricted optimality of linear codes.*

Our converse is based on showing the existence of at least one *special user* who can decode a certain number of messages outside its side information set.

The number of messages decoded by a user is a lower bound on the code length (because it is related to the maximum entropy of the received signal). Therefore, finding the maximum number of messages that a user must be able to decode provides a converse bound on the optimal number of transmissions. The stumbling block in previous approaches, such as the one in [8], was how to find such a special user. The problem of finding the special user can be approached in two ways: 1) constructively finding such a special user for each choice of desired messages, or 2) implicitly proving its existence. In this work we use both methods.

Constructive Method: For the *complement-consecutive* complete–S PICOD(t), which is the complete–S PICOD(t) with $S = [0 : m-t] \backslash [s_{\min} : s_{\max}]$ where $0 < s_{\min} \leqslant s_{\max} < m-t$, we constructively find that the special user, which is the one whose side information set is empty, can always decode $|S| + 2t - 2$ messages.

Combinatorial Method: The constructive method is not amenable for the consecutive complete–S PICOD(t), which is the complete–S PICOD(t) with $S = [s_{\min} : s_{\max}]$ where $0 \leqslant s_{\min} \leqslant s_{\max} \leqslant m-t$, due to the large number of sub-cases / different desired message set assignments that must be considered separately. Therefore for this case we propose a novel combinatorial proof to show the existence of a special user that can decode a certain number of messages. By not only focusing on the desired messages, but on all the messages that a user is eventually able to decode, we consider the messages that a user can eventually know as a *block cover* for this user's side information set; the terminology is borrowed from the combinatorial design structure known as Steiner systems [14]. We argue that the absence of a special user leads to a contradiction to the property of the block cover, and that therefore the special user must exist. This new technique greatly reduces the complexity of the proof compared to the constructive method and enables us to obtain a converse bound for a very general class of complete–S PICOD(t) problems. The keystone of the proof is to show that, for the *critical* complete–S PICOD(t) case with $S = \{s\}$ and $m = 2s+t$, there must exist at least one user who can decode $s+t$ messages. From this, the extension to the consecutive complete–S PICOD(t) follows by enhancing the system to a critical one. By similar reasoning, all complete–S PICOD(t)

where the number of messages is $m \leqslant 5$ can be solved.

The idea of showing the existence of a special user can also be used for the following PICOD(t) problem – for a detailed definition please refers to Section VIII-A. For the case $t = 1$ we show a tight converse for those PICOD(1) with circular-arc network topology hypergraph. For this setting, when a 1-factor does not exist, we show that the code length is at least two by finding a user that can decode two messages.

D. Paper Organization

The rest of the paper is organized as follows: Section II introduces the system model and related definitions; Section III presents the main results of this paper; Sections IV-VII present converse proofs for some complete–S PICOD(t) problems and their optimality; Section VIII shows the optimal information theoretic converse for the PICOD(1) with circular-arc network topology hypergraph; Section IX concludes the paper and discusses future work; some proofs can be found in the Appendix.

E. Notation

Throughout the paper we use capital letters to denote sets, calligraphic letters for family of sets, and lower case letters for elements in a set. The cardinality of the set A is denoted by $|A|$. For integers a_1, a_2 we let $[a_1 : a_2] := \{a_1, a_1+1, \ldots, a_2\}$ for $a_1 \leqslant a_2$ and $[a_2] := [1 : a_2]$ for $a_2 \geqslant 1$. A capital letter as a subscript denotes set of elements whose indices are in the set, i.e., $W_A := \{w_a : w \in W, a \in A\}$. For two sets A and B, $A \backslash B$ is the set that consists all the elements that are in A but not in B. Notations and nomenclature from graph theory will be introduced in Section VIII.

II. System Model

In a PICOD(t) system there is one server, or transmitter, and $n \in \mathbb{N}$ clients, or users; the user set is denoted as $U := \{u_1, u_2, \ldots, u_n\}$. The server is connected to all users via a rate-limited noiseless broadcast channel. There are $m \in \mathbb{N}$ independent and uniformly distributed binary messages of $\kappa \in \mathbb{N}$ bits each; the message set is denoted as $W := \{w_1, w_2, \ldots, w_m\}$. User u_i has a subset of the message set as its side information set $A_i \subset [m]$, $i \in [n]$ [1]. The collection of all side information sets is denoted as $\mathcal{A} := \{A_1, A_2, \ldots, A_n\}$; $\mathcal{A}$ is assumed globally known at the transmitter and all users.

The server broadcasts to the users a codeword of length $\ell\kappa$ bits, which is a function of the message set W and the collection of all side information sets $\mathcal{A}$, i.e., for some function ENC we have

$$x^{\ell\kappa} = \mathsf{ENC}(W, \mathcal{A}). \tag{1}$$

Each user decodes based on the received $x^{\ell\kappa}$, its own side information set, and the collection of all side information sets $\mathcal{A}$; for user $u_j, j \in [n]$, the decoding function is

$$\{\widehat{w}_1^{(j)}, \ldots, \widehat{w}_t^{(j)}\} = \mathsf{DEC}_j(W_{A_j}, x^{\ell\kappa}, \mathcal{A}). \tag{2}$$

[1]Note that if $m - |\cup_{i\in[n]} A_i| \geqslant t$ the problem becomes trivial.

Every user must be able to reliably decode at least t messages not in its side information set, i.e., the decoding error probability for decoding functions $\{\mathsf{DEC}_j, \forall j \in [n]\}$ satisfies

$$\Pr\Big[\exists j \in [n] : \forall \{d_{j,1}, \ldots, d_{j,t}\} \cap A_j = \varnothing, \\ \{\widehat{w}_1^{(j)}, \ldots, \widehat{w}_t^{(j)}\} \neq \{w_{d_{j,1}}, \ldots, w_{d_{j,t}}\}\Big] \leqslant \epsilon_\kappa, \tag{3}$$

for some $\epsilon_\kappa \in (0, 1)$. For a reliable code, $\{\widehat{w}_1^{(j)}, \ldots, \widehat{w}_t^{(j)}\} = \{w_{d_{j,1}}, \ldots, w_{d_{j,t}}\}$ is called the *desired message set* for user u_j, $j \in [n]$, and the indices of the desired messages are denoted as $D_j := \{d_{j,1}, \ldots, d_{j,t}\}$ where $D_j \cap A_j = \varnothing, \forall j \in [n]$. The choice of desired messages for the users is denoted as $\mathcal{D} := \{D_1, D_2, \ldots D_n\}$. The goal is to find the shortest code-length with vanishing-error[2], that is,

$$\ell^* := \inf\{\ell : \exists \text{ a reliable code of length } \ell\kappa \\ \text{such that } \lim_{\kappa\to\infty} \epsilon_\kappa = 0\}. \tag{4}$$

In the following we shall mainly focus on the *complete–S PICOD(t)*, for a given set $S \subseteq [0 : m - t]$. In this system, there are $n := \sum_{s\in S} \binom{m}{s}$ users, where no two users have the same side information set. In other words, all possible users with distinct side information sets that are subsets of size s of the message set, for all $s \in S$, are present in the complete–S PICOD(t).

III. Main Results and Discussion

This section summarizes our main results and comments on their proof techniques, their relationship with past work, and their implications. We start with a simple achievable scheme based on linear codes, in Section III-A. The main contribution of the paper is converse bounds on the optimal code length for the two broad families of PICOD(t): (i) the complete-S PICOD(t), including complement-consecutive S, consecutive S, and their extensions, in Section III-B, and (ii) the PICOD(1) with circular-arc network topology hypergraph, in Section III-C.

A. Achievability

We give next an achievable scheme for the general complete–S PICOD(t) based on linear codes.

Proposition 1 (Achievable Scheme)**.** *Let $\mathcal{S}$ by a partition of S, i.e., $S = \cup_{i\in[|\mathcal{S}|]} S_i$ and $S_i \cap S_j = \varnothing$ for all $i, j \in [|\mathcal{S}|]$ such that $i \neq j$. The optimal code length for the complete–S PICOD(t) with m messages is upper bounded by*

$$\ell^* \leqslant \sum_{i\in[|\mathcal{S}|]} \min\left\{m - \min_{s\in S_i}\{s\}, \max_{s\in S_i}\{s\} + t\right\}. \tag{5}$$

[2]The zero-error setting, that is, where in (3) we impose that $\epsilon_\kappa = 0$ for some κ, is more restrictive than the vanishing-error setting used here. We note that our converse bounds will be derived for the vanishing-error setting, but the achievability bounds under the zero-error setting. We also note that in classical information theory one defines a family of $(2^{\mathsf{n}R}, \mathsf{n}, \epsilon_\mathsf{n})$ codes (indexed by the block-length n) with $2^{\mathsf{n}R}$ codewords, each of length n channel uses, and with probability of error ϵ_n; in the vanishing-error setting one is interested in the largest rate R such that $\lim_{\mathsf{n}\to\infty} \epsilon_\mathsf{n} = 0$. Our setting, as in [1], is the classical information theoretical definition if one identifies $\kappa = \mathsf{n}R$ and $\ell = 1/R$.

By minimizing over all possible partitions $\mathcal{S}$, we have

$$\ell^* \leqslant \min_{\mathcal{S}} \sum_{i\in[|\mathcal{S}|]} \min\left\{m - \min_{s\in S_i}\{s\}, \max_{s\in S_i}\{s\} + t\right\}. \quad (6)$$

The proof is simple and can be deduced from Remark 1.

Remark 1. *Proposition 1 is a generalization of the scheme proposed in [5] whose main idea is as follows. Let $s_{\min}$ and $s_{\max}$ denote the smallest and largest size of the side information sets, respectively. Transmitting $s_{\max}+t$ messages one by one can satisfy all users since each user has at most $s_{\max}$ messages in its side information set. Transmitting $m - s_{\min}$ linearly independent linear combinations of the m messages also satisfies all users, as each user has at least $s_{\min}$ messages in its side information set. Therefore by choosing the best of these two linear codes, we have $\ell^* \leqslant \min\{s_{\max}+t, m-s_{\min}\}$.*

We generalize this idea for the complete–S PICOD(t) by partitioning S into the collection $\mathcal{S}$ and by satisfying the users in each $S_i \in \mathcal{S}$ by using the above scheme. The total code length is the sum of the length of the code used in each partition. Finally, the shortest code length this scheme can achieve is given by searching the best possible partition of S.

B. Converse for some complete–S PICOD(t) problems

We show that for two choices of S the achievability in Proposition 1 is information theoretic optimal.

Theorem 1 (Converse for the *complement-consecutive* complete–S PICOD(t))**.** *For the complete–S PICOD(t) with m messages and $S = [0:m-t]\backslash[s_{\min}:s_{\max}] = [0:s_{\min}-1]\cup[s_{\max}+1:m-t]$ for some $0 < s_{\min} \leqslant s_{\max} < m-t$ (note that the set S includes elements 0 and $m-t$), the optimal code length is*

$$\begin{aligned}\ell^* &= \min\{m, m+t+s_{\min}-s_{\max}-2\}\\ &= \min\{m, |S|+2t-2\}. \end{aligned} \quad (7)$$

The proof of Theorem 1 can be found in Section IV.

Theorem 2 (Converse for the *consecutive* complete–S PICOD(t))**.** *For the complete–S PICOD(t) with m messages and $S = [s_{\min}:s_{\max}]$ for some $0 \leqslant s_{\min} \leqslant s_{\max} \leqslant m-t$ (i.e., S contains consecutive integers, from $s_{\min}$ to $s_{\max}$) the optimal code length is*

$$\ell^* = \min\{s_{\max}+t, m-s_{\min}\}. \quad (8)$$

The proof of Theorem 2 is broken down in several pieces. The proof for the *critical case*, where $m = 2s+t$ and $S = \{s\}$, can be found in Section V, while the general proof is presented in Section VI.

Remark 2. *Theorems 1 and 2 show that the simple achievable scheme in Proposition 1 is information theoretically optimal for a class of PICOD(t). Specifically, the consecutive complete–S PICOD(t) is the oblivious PICOD(t) studied in [5]. Our Theorem 2 provides a tight information theoretic converse for the achievability proposed in [5].*

The basic idea in the proof of Theorem 1 is to prove the existence of a user who can decode $|S|$ messages by a method referred to as layer counting. *We partition all users in the complete–S PICOD(t) into $|S|$ layers. Each layer contains the users with the same size of the side information set. A layer is said to be "lower" than another if the size of the side information set of the users is smaller. The intuition is that a user in a lower layer, after having decoded its desired messages, can mimic users in higher layers and thus decode also the desired messages of those higher layer users.*

In the complement-consecutive complete–S PICOD(t), where $S = [0:s_{\min}-1]\cup[s_{\max}+1:m-t]$ for some $0 < s_{\min} \leqslant s_{\max} < m-t$, we show the user in the lowest layer (with empty side information set) can mimic a user in each higher layers and eventually decodes $|S|+2t-2$ messages.

However, this layer counting converse is not tight in general, as explained in Remark 6 for the complete–S PICOD(1) with $S = [1:q]$ or $S = [q:m-2]$ for some $2 \leqslant q \leqslant m-2$. To improve on the layer counting converse, we propose a novel converse technique in Theorem 2 for the consecutive *complete–S PICOD(t), where $S = [s_{\min}:s_{\max}]$ for some $0 \leqslant s_{\min} \leqslant s_{\max} \leqslant m-t$. The* critical case *for this proof is the complete–S PICOD(t) for*

$$m = 2s+t \text{ messages and } S = \{s\} \quad \textit{(critical case)}. \quad (9)$$

In Section V Proposition 6, we show that for this critical case, regardless of the choice of desired messages and valid code, there always exists at least one user who can decode $s+t$ messages. While the proof of Theorem 1 is constructive, that is, we explicitly identify the user who can always decode $|S|+2t-2$ messages (the one with empty side information set), the proof of Proposition 6 is not. The problem with a constructive argument for the critical case is that, for any specific user, there exists an information theoretic optimal choice of desired messages and a corresponding valid code such that this user can decode only its desired t messages and no more. In other words, showing that a certain user can always decode more than t messages is impossible. Therefore, in the proof of Proposition 6, we propose a combinatorial method to show the existence of at a least a user with some desired property, namely, the ability to decode a certain number of messages. The new method involves the Maximum Acyclic Induced Subgraph (MAIS) converse idea for the classic IC [3], as well as a combinatorial design technique inspired by Steiner systems [14], which we shall refer to as block cover. *The existence proof does not indicate which user has the desired property, but only shows its existence regardless of the choice of desired messages at the users.*

Theorem 2 can be further extended to cover other complete–S PICOD(t). We have the following results.

Proposition 2 (Not a complete–S system, but all users are below the critical case users in the layer representation)**.** *For the complete–S PICOD(t) with m messages and $s_{\max} := \max_{s\in S}\{s\} \leqslant \lfloor\frac{m-t}{2}\rfloor$, the optimal code length is $\ell^* = s_{\max}+t$.*

The proof can be found in Section VII.

Proposition 3 (Not a complete–S system, but all users are above the critical case users in the layer representation)**.** *For the complete–S PICOD(t) with m messages and*

$s_{\min} := \min_{s\in S}\{s\} \geqslant \lceil \frac{m-t}{2} \rceil$, *the optimal code length is* $\ell^* = m - s_{\min}$.

The proofs can be found in Section VII.

Proposition 4 (Not a complete–S system, but all users in a band around the critical case users are present in the layer representation)**.** *For the complete–S PICOD*(t) *with* m *messages, let*

$$\delta := \min\left\{ s_{\max} - \lceil \frac{m-t}{2} \rceil, \lfloor \frac{m-t}{2} \rfloor - s_{\min} \right\}, \tag{10}$$

where $s_{\max} := \max_{s\in S}\{s\}$ *and* $s_{\min} := \min_{s\in S}\{s\}$. *If* $\left[\lfloor \frac{m-t}{2} \rfloor - \delta : \lceil \frac{m-t}{2} \rceil + \delta\right] \subseteq S$ *then the optimal code length is* $\ell^* = \min\{s_{\max} + t, m - s_{\min}\}$.

The proof can be found in Section VII.

Remark 3. *Propositions 2, 3 and 4 show an interesting fact: for these settings the only relevant layers in the layer representation are the ones closest to the "critical" middle layer* $\frac{m-t}{2}$*, or the layers in a band* $\left[\lfloor \frac{m-t}{2} \rfloor - \delta : \lceil \frac{m-t}{2} \rceil + \delta\right]$ *around the "critical" middle layer. The optimal code for the users in these layers satisfies all the remaining users.*

Finally, for those PICOD(t) problems with $m \leqslant 5$ messages that are not covered by Propositions 2, 3, 4 and Theorem 1, we have the following:

Proposition 5. *For all complete–S PICOD*(t) *with* $m \leqslant 5$ *and non-empty* $S \subseteq [0 : m-1]$*, the achievable scheme in Proposition 1 is information theoretic optimal.*

The proof can be found in Section VII.

Remark 4. *Proposition 5 is proved by checking one by one all complete–S PICOD*(t) *problems with* $m \leqslant 5$ *messages not covered by previous results. It may be possible to go beyond five messages, but unfortunately we have not been able to find a systematic way to prove the converse for general* m*.*

C. Converse for the PICOD(1) *with circular-arc network topology hypergraph*

The reader can find a refresher on graph theory terminology in Section VIII-A. The critical complete–$\{s\}$ PICOD(t) we solved has a network topology hypergraph which is the dual hypergraph of the complete $(m-s)$–uniform hypergraph. Here we solve the PICOD(1) whose network topology hypergraph is a special hypergraph, namely, a circular-arc hypergraph.

Theorem 3. *For a PICOD*(1) *with* m *messages and with circular-arc network topology hypergraph, the optimal code length satisfies* $\ell^* \leqslant 2$*. In particular, the optimal number of transmissions is* $\ell^* = 2$ *unless the network topology hypergraph is a* 1*-factor hypergraph.*

The proof can be found in Section VIII.

Remark 5. *The achievability part of Theorem 3 is based on the following property of a circular-arc hypergraph: if two vertices belong to an edge, then all vertices (cyclic) between these two vertices must belong to the same edge. The converse part of Theorem 3, which is in Proposition 8, is proved by showing that there exists a user that can decode one more message other than its desired message if a* 1*-factor does not exist. By showing the existence of such a user, regardless of the choices of desired messages and code sent by the transmitter, we obtain a tight lower bound on the optimal code length.*

The proofs of the converse results summarized in this section will be given in the following sections.

IV. Layer Counting Converse: Proof of Theorem 1

Recall that the complete–S PICOD(t), for a given set $S \subseteq [0 : m-t]$, comprises $n = \sum_{s\in S} \binom{m}{s}$ users where the side information sets are all possible distinct subsets of size s of m messages, for all $s \in S$. The proof of Theorem 1 relies on idea of *decoding chain*, which gives a high level explanation of the proof of the following lemma (see discussion after the proof), namely, the number of messages decoded along this chain provides a converse on ℓ^*.

Lemma 1. *In a PICOD*(t) *with* m *messages and* n *users, for any ordering of the users (i.e., up to relabeling the users) we have*

$$\ell^* \geqslant \sum_{i=1}^{n} \left| D_i \backslash \cup_{j=1}^{i-1} (A_j \cup D_j) \right|. \tag{11}$$

Proof of Lemma 1: Since we have a working system, all users are satisfied by the transmission of $x^{\ell\kappa}$ of length ℓ. For user u_1 we have

$$H\left(W_{D_1} | x^{\ell\kappa}, W_{A_1}\right) \leqslant \kappa\epsilon_\kappa, \tag{12}$$

where $\lim_{\kappa\to\infty} \epsilon_\kappa = 0$ by Fano's inequality. Similarly, for user u_2 we have

$$H\left(W_{D_2} | x^{\ell\kappa}, W_{A_2}\right) \leqslant \kappa\epsilon_\kappa. \tag{13}$$

Therefore we have

$$\begin{aligned}
&H\left(W_{D_1}, W_{D_2} | x^{\ell\kappa}, W_{A_1}, W_{A_2\backslash D_1}\right) \\
&= H\left(W_{D_1} | x^{\ell\kappa}, W_{A_1}, W_{A_2\backslash D_1}\right) \\
&+ H\left(W_{D_2} | x^{\ell\kappa}, W_{A_1}, W_{A_2\backslash D_1}, W_{D_1}\right) \\
&= H\left(W_{D_1} | x^{\ell\kappa}, W_{A_1}, W_{A_2\backslash D_1}\right) \\
&+ H\left(W_{D_2} | x^{\ell\kappa}, W_{A_2}, W_{A_1\cup D_1}\right) \\
&\leqslant H\left(W_{D_1} | x^{\ell\kappa}, W_{A_1}\right) + H\left(W_{D_2\backslash(A_1\cup D_1)} | x^{\ell\kappa}, W_{A_2}\right) \\
&\leqslant 2\kappa\epsilon_\kappa.
\end{aligned}$$

By continuing with the same reasoning, we get

$$H\left(W_{\cup_{i=1}^n D_i} | x^{\ell\kappa}, W_{\cup_{i=1}^n (A_i\backslash\cup_{j=1}^{i-1} D_j)}\right) \leqslant n\kappa\epsilon_\kappa. \tag{14}$$

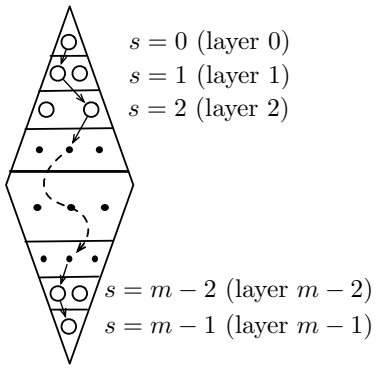

Fig. 1: Layer representation of the complete–$[0 : m-1]$ PICOD(1) problem.

Since the messages are independent and uniformly distributed with entropy κ bits, and since the code is binary, we conclude

$$\begin{aligned}
&\sum_{i=1}^{n} \left|D_i \backslash \cup_{j=1}^{i-1} (A_j \cup D_j)\right| \kappa \\
&= \left|\cup_{i=1}^{n} \left(D_i \backslash \cup_{j=1}^{i-1} (A_j \cup D_j)\right)\right| \kappa \\
&= H\left(W_{\cup_{i=1}^{n}\left(D_i \backslash \cup_{j=1}^{i-1}(A_j \cup D_j)\right)}\right) \\
&= H\left(W_{\cup_{i=1}^{n}\left(D_i \backslash \cup_{j=1}^{i-1}(A_j \cup D_j)\right)} | W_{\cup_{i=1}^{n}(A_i \backslash \cup_{j=1}^{i-1} D_j)}\right) \\
&\leqslant I\left(W_{\cup_{i=1}^{n}\left(D_i \backslash \cup_{j=1}^{i-1}(A_j \cup D_j)\right)}; x^{\ell\kappa} | W_{\cup_{i=1}^{n}(A_i \backslash \cup_{j=1}^{i-1} D_j)}\right) \\
&+ n\kappa\epsilon_\kappa \\
&\leqslant H\left(x^{\ell\kappa} | W_{\cup_{i=1}^{n}(A_i \backslash \cup_{j=1}^{i-1} D_j)}\right) + n\kappa\epsilon_\kappa \\
&\leqslant H(x^{\ell\kappa}) + n\kappa\epsilon_\kappa \\
&\leqslant \ell\kappa + n\kappa\epsilon_\kappa,
\end{aligned}$$

which implies that

$$\ell \geqslant \sum_{i=1}^{n} \left|D_i \backslash \cup_{j=1}^{i-1} (A_j \cup D_j)\right|, \tag{15}$$

for constant n, sufficiently large κ, and any valid codes. Therefore the bound in (15) must hold for the optimal code length as well, thus proving (11). ∎

The sequence of users $u_1, u_2, \ldots, u_n$ in Lemma 1 is the *decoding chain* mentioned at the beginning of this section. In fact, the converse in Lemma 1 can also be thought of as the *acyclic induced subgraph* converse for the all unicast IC problem [3], where each user desires multiple messages, as opposed to a single message. The users with $\left|D_i \backslash \cup_{j=1}^{i-1} (A_j \cup D_j)\right| \neq 0$ form an acyclic induced subgraph in the graph representation of the IC. Therefore, in Lemma 1 the value of $\left|\cup_{i=1}^{n} \left(D_i \backslash \cup_{j=1}^{i-1} (A_j \cup D_j)\right)\right|$ depends on the choice of the order for the users, that is, we can relabel the users in order to find the tighest bound provided by Lemma 1. Finding such an order for Lemma 1 illustrates the intuition for the converse proof of Theorem 1: finding the user that can decode the largest number of messages.

To illustrate the method of finding the user who can decode the largest number of messages, we introduce the *layer representation* of the complete–S PICOD(t). As an example, the layer representation for the complete–$[0 : m-1]$ PICOD(1) problem is given in Fig. 1. In Fig. 1, all the users with the same size of the side information set are said to form a layer, and there are in total m layers; the i-th layer contains the users whose side information set has size $i \in [0 : m-1]$, and the number of users in the i-th layer is $\binom{m}{i}$. The key observation is that, in a working system, a user u_i in i-th layer can decode a message w_{d_i} it does not have in its side information set A_i. After that, user u_i is equivalent to a user u_{i+1} in the $(i+1)$-th layer whose side information is $A_{i+1} = A_i \cup \{d_i\}$. User u_i will thus be able to decode the message $w_{d_{i+1}}$ that is desired by user $u_{d_{i+1}}$, in addition to its own desired message w_{d_i}. But now user u_i will have $A_{i+2} = A_i \cup \{w_{d_i}, w_{d_{i+1}}\}$, which is the side information of a user u_{i+2} in the $(i+2)$-th layer. By continuing with the same reasoning, user u_i will be able to mimic one user per layer until the last layer. We apply this argument to the user in the 0-th layer (there is only one such user). We see that the user in the 0-th layer is able to decode one message per layer without loss of optimality, that is, the user in the 0-th layer decodes m messages in total. This provides a decoding chain of length m. In this decoding chain each user's side information set and the desired message set form the side information set of the next user. By having such a decoding chain, we can use Lemma 1 to show that $\ell^* \geqslant m$ for the complete–$[0 : m-1]$ PICOD(1) problem in Fig. 1. We use this observation, and similar ones, in the following to provide a lower bound on ℓ^* in terms of number of messages a user can decode, which is the main idea in all our converse proofs.

The proof of Theorem 1 directly follows this idea of counting the layers in a layer representation of a decoding chain. The key for the proof is the fact that each layer in the layer representation for the complement-consecutive complete–S, where $S = [0 : m-t] \backslash [s_{\min} : s_{\max}]$, contains all users with side information set of the same size. After the user has decoded its desired message(s), we can map this user to another user in a higher layer. Such a mapping forms a decoding chain, starting from the user in the 0-th layer, and provides a lower bound on ℓ^*.

We are now ready to prove Theorem 1 for the complement complete–S with m messages and $S = [0 : m-t] \backslash [s_{\min} : s_{\max}]$.

Proof of Theorem 1: Consider the PICOD(t) where $S = [0 : m-t] \backslash [s_{\min} : s_{\max}] = [0 : s_{\min} - 1] \cup [s_{\max} + 1 : m-t]$ for some $0 < s_{\min} \leqslant s_{\max} < m - t$. We aim to find the decoding chain that has the largest number of messages/users along the chain. In each layer of the layer representation we find a user for the decoding chain. Therefore the chain contains $|S|$ users which we shall indicate as $\{u_1, u_2, \ldots, u_{|S|}\}$, where

$$|S| = s_{\min} + m - t - s_{\max}. \tag{16}$$

We first find $s_{\min}$ users, one user per layer for the layers indexed by $[0 : s_{\min} - 1]$, as done in the example in Fig. 1. Then, the user $u_j|_{j=s_{\min}+1}$ is found in layer $s_{\max} + 1$ such that $|A_j|_{j=s_{\min}+1}| = s_{\max} + 1$; we want $A_{s_{\min}+1} \supseteq A_{s_{\min}}$ so user $u_{s_{\min}+1}$ can be mimicked by user $u_{s_{\min}}$; we want $|D_{s_{\min}} \cap A_{s_{\min}+1}|$ to be as large as possible so the number of messages decoded in the decoding chain can be maximized;

details on how this is done will be given next. Finally, we find other $|S| - s_{\min} - 1 = m - t - s_{\max} - 1$ users to complete the decoding chain, one user per layer for the layers indexed by $[s_{\max} + 2 : m - t]$, as done in the example in Fig. 1.

Assume all users are satisfied by the transmission of $x^{\ell\kappa}$. Let u_1 be the user with empty side information set, i.e., $A_1 = \varnothing$. Since all users are satisfied, u_1 can decode at least one message not in its side information set; denote the index of such a message as $d_{1,1}$. Layer 1 contains the users with side information set of size 1. There exists a user in layer 1, say u_2, with side information $A_2 = A_1 \cup \{d_{1,1}\} = \{d_{1,1}\}$ and desired message $d_{2,1} \notin A_2$. By continuing with this reasoning we can find users up to user $u_{s_{\min}}$: user $u_{s_{\min}}$ has side information set $A_{s_{\min}} = A_{s_{\min}-1} \cup \{d_{s_{\min}-1,1}\} = \{d_{1,1}, \ldots, d_{s_{\min}-1,1}\}$ and desires message $d_{s_{\min},1} \notin A_{s_{\min}}$.

We would be tempted to say that the $(s_{\min} + 1)$th user in the decoding chain, which is denoted as $u_{s_{\min}+1}$, should have side information set $A_{s_{\min}} \cup D_{s_{\min}}$ and be in layer $s_{\min}+t-1$. However $|A_{s_{\min}} \cup D_{s_{\min}}| = |A_{s_{\min}}| + |D_{s_{\min}}| = s_{\min} - 1 + t$ may be strictly less than the size of the side information of the next layer of users present in the system; if so, a user with side information set $A_{s_{\min}} \cup D_{s_{\min}}$ does not exist. For this reason, we choose user $u_{s_{\min}+1}$ in layer $s_{\max}+1$ as follows: if $s_{\min}-1+t < s_{\max}+1$, we choose as $u_{s_{\min}+1}$ any user in layer $s_{\max}+1$ that satisfies $A_{s_{\min}+1} \supset A_{s_{\min}} \cup D_{s_{\min}}$, i.e., the user that $u_{s_{\min}}$ can mimic by providing the messages indexed by $A_{s_{\min}+1} \backslash (A_{s_{\min}} \cup D_{s_{\min}})$ as genie side information; otherwise we choose $u_{s_{\min}+1}$ to be the user with $A_{s_{\min}+1} \subseteq A_{s_{\min}} \cup D_{s_{\min}}$, i.e., the user that $u_{s_{\min}}$ can mimic. With this choice for user $u_{s_{\min}+1}$ we insure that user $u_{s_{\min}}$ has all messages in the side information set of the user $u_{s_{\min}+1}$.

From this point onwards, the next users in the decoding chain can again be chosen such that $A_j = A_{j-1} \cup \{d_{j-1,1}\},\ j \in [s_{\min} + 2 : |S|]$.

Note that in the decoding chain we have $A_j = A_{j-1} \cup \{d_{j-1,1}\}$ for $j \in [|S|] \backslash \{1, s_{\min} + 1\}$, $A_1 = \varnothing$, and $A_{s_{\min}+1} \supset A_{s_{\min}}$. These users satisfy $A_j \supset A_{j-1}$, for all $j \in [2 : |S|]$. Therefore we have $\left|D_i \backslash \cup_{j=1}^{i-1} (A_j \cup D_j)\right| \geqslant 1$ for all $i \in [|S|] \backslash \{1, s_{\min} + 1\}$, $|D_i|_{i=1} = t$, and $\left|D_i \backslash \cup_{j=1}^{i-1} (A_j \cup D_j)\right|_{i=s_{\min}+1} = \min\{t, s_{\max}+1+t-(s_{\min}-1+t)\} = \min\{t, s_{\max} - s_{\min} + 2\}$. Therefore, by Lemma 1 we have

$$\begin{aligned}
\ell^* &\geqslant \sum_{i \in [|S|]} \left|D_i \backslash \cup_{j=1}^{i-1} (A_j \cup D_j)\right| \\
&\geqslant t + (s_{\min} - 1) + \min\{t, s_{\max} - s_{\min} + 2\} \\
&\quad + (m - t - s_{\max} - 1) \\
&\geqslant m + s_{\min} - s_{\max} - 2 + \min\{t, s_{\max} - s_{\min} + 2\} \\
&= \min\{m + t + s_{\min} - s_{\max} - 2, m\}.
\end{aligned}$$

The value $\ell^* = \min\{m + t + s_{\min} - s_{\max} - 2, m\}$ can be achieved as follows. By the scheme in Proposition 1 with partition $S = S_1 \cup S_2$, with $S_1 := [0 : s_{\min} - 1]$ and $S_2 := [s_{\max} + 1 : m - t]$, all users in group S_1 are satisfied with $(s_{\min} - 1) + t$ transmissions, and all users in group S_2 are satisfied with $m - (s_{\max} + 1)$ transmissions; therefore, we have a code of length $m + s_{\min} + t - s_{\max} - 2$. Also, we can always transmit all m messages one by one, resulting in a code of length m. Therefore, we can achieve the lower bound by using the code among the above two with the shortest length.

This concludes the proof of Theorem 1. ■

Remark 6. *The proof of Theorem 1 constructively builds a decoding chain. The decoding chain starts from the user in the lowest layer. The next user in the chain is chosen in the next layer, based on the side information and desired message of the previous one. The chain ends at the highest layer. However, this construction, where each layer contributes at most one user to the decoding chain, is not always tight.*

As shown in [8], for the complete–S PICOD(1) where $S = [1 : q]$ or $S = [q : m - 2]$, $1 \leqslant q \leqslant m - 2$, the optimal code length is $\ell^* = |S| + 1$. In other words, there exists a decoding chain which includes two users with the same size of side information, where one of the users can mimic the other one.

The proof in [8] is a case-by-case reasoning, where the different cases are for different choices of desired messages of the users. For the complete–S PICOD(1) for general $S = [s_{\min} : s_{\max}]$, the number of cases becomes too large to be tractable. Thus a method that does not relay on a case-by-case study becomes necessary. This is what we are going to do in the next section. The two cases considered in [8] are special cases of Theorem 2 proved next.

V. Critical Case: complete–$\{s\}$ PICOD(t) with $m = 2s + t$ messages

To overcome the limitation of the case-by-case reasoning highlighted in Remark 6, we shall turn to an *existence proof* technique for Theorem 2. Loosely speaking, when dealing with general consecutive complete–S PICOD(t) with $S = [s_{\min} : s_{\max}]$, we treat all users and all the various desired message assignments at once. Before we prove Theorem 2 in full generality, we consider the critical case in (9). We shall see that all other consecutive complete–S cases can be deduced from the critical one. Therefore, this section contains the proof for the following key result:

Proposition 6. *(The critical case) For the complete–$\{s\}$ PICOD(t) with $m = 2s + t$ messages, the optimal code length is $\ell^* = s + t$. Specifically, given a valid code, there always exists a user that can decode $\ell^* = s + t$ messages.*

As for the layer counting converse used in Theorem 1, we shall show that under the assumption that all users can decode at least one message outside their side information set, there must exists a user that can mimic other users and decodes $\ell^* = s + t$ messages regardless of the desired messages of all the users. Note that in the complete–S PICOD(t) where $|S| = 1$, only one layer exists in the layer representation. Thus by the constructive method in Theorem 1, we only obtain the trivial bound $\ell^* \geqslant 1$. However, we do need to find the specific user that can decode $\ell^* = s + t$ messages, but only show its existence. So we turn to an existence proof, which is largely based on combinatorics ideas. Specifically, for all possible desired message set assignments for the users, given a valid code that satisfies all users, we show that there exists a user that can decode $\ell^* = s + t$ messages. We start by

introducing next the two main ingredients needed in the proof of Proposition 6.

A. Proposition 6: Converse Main Ingredient 1: Block Cover

So far we used the idea of decoding chain to show that a user can decode more than its set of desired messages. The decoding chain depends on the choice of desired messages at the users. Once the desired messages change, the decoding chain may change as well. Here we are only interested in the existence of such a decoding chain of a given length. In other words, we show the existence of a decoding chain of a certain length regardless of the choice of desired messages at the users. We start with a simple example to showcase a problem we faced when considering different message assignments.

Example 1. *Consider the complete–$\{1\}$ PICOD(1), i.e., $s = t = 1$, with $m = 2s+1 = 3$ messages for which $\ell^\star = s+1 = 2$ is the smallest number of transmissions needed to satisfy all the $n = \binom{m}{s} = 3$ users. Say that u_1 knows $A_1 = \{1\}$ and desires $d_1 = 2$; u_2 knows $A_2 = \{2\}$ and desires $d_2 = 1$; and u_3 knows $A_3 = \{3\}$ and desires $d_3 = 1$. By sending w_1, users u_2 and u_3 are satisfied; by sending w_2, user u_1 is satisfied. By the decoding chain argument, user u_3 is able to mimic u_1 (because he decodes the message that is the side information set of user u_1) and therefore can also decode w_2; on the contrary, users u_2 and u_3 can not decode any more messages other than the desired one. However, another choice of desired messages can be $d_1 = 3, d_2 = 1, d_3 = 1$; with this, users u_1 and user u_3 can only decode their desired messages while user u_2 can mimic user u_1 thus is able to decode two messages.*

As Example 1 shows for the case $t = 1$, for a specific user, there is always an optimal choice of desired messages such that this user cannot decode any message other the desired one. However, we also note that for any choice of desired messages, there always exists a user that can decode two messages. In the critical case setting, we shall prove that regardless of the choice of desired messages, there always exists a user who can decode $s+t$ messages. Since there are $\binom{s+t}{t}^{\binom{2s+t}{s}}$ (doubly exponential in s) possible choices of desired messages, finding explicitly such a user for every case is intractable. Therefore, our converse shows the existence of such a user. The main idea of the existence proof is as follows.

Instead of checking all possible different choices of desired message sets at the users, we reason on the size of the decoding chain for that user. By assumption, every user can decode t messages outside its side information set. Some users may be able to decode more messages because they can mimic other users. After receiving a valid code, we aim to show that every user eventually knows at least $s+t$ messages, including the s messages in its side information set and the (at least) t decoded ones. Say that user u_j, with side information A_j, eventually can decode the messages indexed by $B_j \supseteq D_j$. One can think of the set $C_j := A_j \cup B_j$ as a *block* that *covers* the side information set A_j, by which we mean that the set C_j is a proper superset of A_j. User u_j can also mimic any users u_k whose side information set satisfies $A_k \subset C_j$. Therefore the desired message sets for all users u_k whose side information $A_k \subset C_j$ satisfy $D_k \subset C_j$. For any set $U' \subseteq [n]$ of the users we can find a collection $\mathcal{C}$ such that, for every side information set $A_j, j \in U'$, there is a cover $C_j \in \mathcal{C}$ such that $C_j = A_j \cup B_j$, where B_j is the largest set of the messages that user u_j can decode. By this definition, this *block cover* / collection $\mathcal{C}$ satisfies the following properties:

1) [BlockCover-P1] For every s-element subset of $[m]$, there exists at least one $C \in \mathcal{C}$ that contains this subset.
2) [BlockCover-P2] $s < |C| \leqslant m$ for all $C \in \mathcal{C}$.
3) [BlockCover-P3] For all $P \subseteq [|\mathcal{C}|]$, we have $|\cap_{j\in P} C_j| \notin [s : s+t-1]$.

Proof of BlockCover Properties: Properties BlockCover-P1 and BlockCover-P2 follow by the definition of block cover, while property BlockCover-P3 holds because if we have $|\cap_{j\in P} C_j| \in [s : s+t-1]$ for some $P \subseteq [|\mathcal{C}|]$, we can have a user with side information set $A' \subseteq \cap_{j\in P} C_j$ with corresponding decoding set D' and this leads to the following contradiction. By definition of intersection $A' \subset C_j,\ \forall j \in P$; but also by definition of block cover $D' \subset C_j,\ \forall j \in P$; thus $A' \cup D' \subseteq C_j,\ \forall j \in P$, which implies $|\cap_{j\in P} C_j| \geqslant |A' \cup D'| = |A'| + |D'| \geqslant s+t$ that contradicts the starting assumption $|\cap_{j\in P} C_j| \in [s : s+t-1]$. ∎

This block cover idea was inspired by the so-called *generalized Steiner systems* in combinatorial design [14]. An $\mathcal{S}(s, *, m)$ generalized Steiner system consists of blocks / sets such that each subset of size s from the ground set of size m is covered exactly once. In a critical PICOD(t) setting, the collection of blocks $\mathcal{C}$ also covers all s-element subsets of $[m]$ (i.e., all users' side information sets). But our problem is not exactly a generalized Steiner system because an s-element subset may be contained in more than one block as long as it is not an exact intersection of the blocks–see Property BlockCover-P3. Therefore, our block cover can be seen as a relaxed generalized Steiner system. To the best of our knowledge no results are available for this specific relaxed generalized Steiner system.

For the critical case we aim to show that there is a user who can decode $s+t$ messages (as in Example 1). We argue it by contradiction. Assume no user can decode $s+t$ messages, that is, every user can decode at least t and at most $s+t-1$ messages by mimicking other users. In terms of block cover, this indicates that we can have a block cover $\mathcal{C}$ with $\max_{C\in\mathcal{C}}\{|C|\} \leqslant (s+t-1)+s < m = 2s+t$. Our argument of showing that there always exists a user that can decode $t+s$ messages for the critical case is equivalent to showing that a block cover with size at most $2s+t-1$ cannot exist. Our combinatorial proof shows that the existence of a choice of desired messages such that $t+s \leqslant |C_j| \leqslant 2s+t-1, \forall j \in [|\mathcal{C}|]$ leads to the existence of a user that can decode $t+s$ messages, thus $\max|C_j| = 2s+t$, which is a contradiction. Therefore must exists a user whose block cover has size $m = 2s+t$.

B. Proposition 6: Converse Main Ingredient 2: Maximum Acyclic Induced Subgraph (MAIS) Bound

Recall that for a PICOD(t), each user chooses t desired messages outside its side information set. The collection of the desired message sets for all the users users is denoted as

$\mathcal{D} = \{D_1, \ldots, D_n\}$, where $n = \binom{2s+t}{s}$. Once D is chosen, the PICOD(t) reduces to a *multi-cast IC* where each user requests t messages; we can make one user to be t users with the same side information sets but each with a distinct single desired message; the multi-cast IC with n users becomes a multi-cast IC with tn users, each requesting one message.

Similarly to the classic all-unicast IC, we can represent in a directed graph / digraph the side information sets and the desired messages of a multi-cast IC where each user desires a single message [3]. Pick a subset $U \subseteq [tn]$ of users who desire different messages and create a digraph $\mathsf{G}(U)$ as follows. The vertices $V(\mathsf{G}) \subseteq W$ represent the desired messages by the users in U. A directed arc $(w_i, w_j) \in E(\mathsf{G})$ exists if and only if the user who desires w_i has w_j in its side information set. G is called *acyclic* if it does not contain a directed cycle. The *size* of G is the number of the vertices in it, i.e. $|V(\mathsf{G})| = |U|$. For the all-unicast IC, the maximum size of U such that the corresponding digraph $\mathsf{G}(U)$ is acyclic serves as a converse bound on the optimal code length. This converse is known as *maximum acyclic induced subgraph* (MAIS) bound [3].

For the PICOD(t), a similar MAIS bound can be found, which is the maximum size of the acyclic digraph $\mathsf{G}(U)$ created by the choice of users $U \subseteq [tn]$ such that they all desire different messages. Since MAIS depends on the desired message set $\mathcal{D}$, we denote its size as $|\text{MAIS}(\mathcal{D})|$. Thus, for the PICOD$(t)$ as for multi-cast IC, the size of MAIS is a converse bound on ℓ [3], namely, $\ell \geqslant |\text{MAIS}(\mathcal{D})|$.

Finding the MAIS for the all-unicast IC is known to be an NP-hard problem [6] in general. Finding the MAIS for the multi-cast IC appears to be more difficult since one needs to check every possible choice of users with distinct desired messages. Finding the MAIS for the PICOD(t) problem seems even more complicated since each choice of $\mathcal{D}$ in the PICOD(t) corresponds to a multi-cast IC, and in addition one needs to find the best $\mathcal{D}$ in terms of code length. Therefore, finding the MAIS for the PICOD(t) by solving all possible all-unicast IC problems appears intractable. Therefore, our existence proof does not find the exact MAIS for the PICOD(t), but only bounds on its size, i.e., $\max_{\mathcal{D}} |\text{MAIS}(\mathcal{D})|$. Towards this goal, we have the following properties:

1) [MAIS-P1] for the critical complete–$\{s\}$ PICOD(t) with $m = 2s + t$ messages, $|\text{MAIS}(\mathcal{D})| = s + t$ for certain $\mathcal{D}$ if and only if there exists a user who decodes $s + t$ messages by mimicking other users.
2) [MAIS-P2] for the critical complete–$\{s\}$ PICOD(t) with $m = 2s + t$ messages, if there exists a $\mathcal{D}$ such that $|\text{MAIS}(\mathcal{D})| \leqslant s + t - 1$, there exists a $\mathcal{D}'$ where $|\text{MAIS}(\mathcal{D}')| = s + t - 1$.

Proof of Property MAIS-P1: On the one hand, if $|\text{MAIS}(\mathcal{D})| = s + t$, there are $s + t$ users who desire different messages. These users form an acyclic induced subgraph. We can obtain a decoding chain from the acyclic induced subgraph, in which the first user has side information of all s messages that are not desired by these $s + t$ users. Since there are in total $m = 2s + t$ messages in the system, the first user, by decoding its desired message, can mimic all the other users and eventually decode $s + t$ messages.

On the other hand, if there is one user who can decode $s+t$ messages, there are $s + t - 1$ users that can be mimicked by it with different desired messages. These $s + t$ users form an acyclic induced subgraph of size $s + t$. Since all users have side information set of size s, $|\text{MAIS}(D)| \leqslant s + t$. We then have $|\text{MAIS}(D)| = s + t$. ∎

Proof of Property MAIS-P2: We prove the claim by showing that for a choice of desired messages $\mathcal{D}$ that has a $|\text{MAIS}(\mathcal{D})| = a$ for some integer $a < s + t$, we can always find another choice of desired messages $\mathcal{D}'$ such that $|\text{MAIS}(\mathcal{D}')| = a + 1$.

Assume there exists a $\mathcal{D}$ such that, for some integer $a < s + t$, satisfies $|\text{MAIS}(\mathcal{D})| = a$. For this $\mathcal{D}$, the PICOD(t) can be seem as a unicast IC with $[tn]$ users, whose graph representation has an induced acyclic subgraph of size a and all induced subgraphs of size strictly larger than a are cyclic. Without loss of generality, let $\{u_1, \ldots, u_a\}$ be the set of users that form this MAIS who have desired messages $\{w_{d_1}, \ldots, w_{d_a}\} = [a]$. By the definition of MAIS, any user with side information $A \subseteq [a + 1 : m]$ must have desired message $d \in [a]$; this is so because any user with $A \subseteq [a+1 : m]$ and $d \in [a+1 : m]$ can be added to the users $u_1, \ldots, u_a$ to form an acyclic subgraph of size $a + 1$, which would contradict to the assumption that $|\text{MAIS}| = a$.

Based on $\mathcal{D}$ we construct $\mathcal{D}'$ such that $|\text{MAIS}(\mathcal{D}')| = a+1$ as follows. Choose a user u' with side information $A' \subseteq [a + 1 : m]$ and change its desired message to $d' \in [a+1 : m] \backslash A'$ (it was $d' \in [a]$ in $\mathcal{D}$). Since $a < s+t$ we have $|[a+1 : m]| \geqslant s+1$ and $|[a + 1 : m] \backslash A'| \geqslant 1$, thus such a user u' and its desired message d' can be found. Moreover, by construction the users in $\{u_1, \ldots, u_a, u'\}$ form an acyclic subgraph of size $a + 1$.

Next, we show that any induced subgraph of size strictly larger than $a + 1$ in the IC represented by $\mathcal{D}'$ is cyclic. This can be seen as follows. Note that from $\mathcal{D}$ to $\mathcal{D}'$ only the desired message of u' was changed, therefore any induced subgraph in the IC represented by $\mathcal{D}'$ that does not have u' also exists in the IC represented by $\mathcal{D}$. For any induced subgraph in the IC represented by $\mathcal{D}'$ with size strictly larger than $a + 1$, if it does not contain u', this induced subgraph exists in the IC represented by $\mathcal{D}$. By the condition $|\text{MAIS}(\mathcal{D})| = a$ we know that this subgraph is cyclic. If the induced subgraph contains u', remove u' so as to obtain an induced subgraph of size strictly larger than a. This newly obtained subgraph exists in the IC represented by $\mathcal{D}$. Similarly the subgraph is cyclic by $|\text{MAIS}(\mathcal{D})| = a$ thus the original subgraph which contains u' is also cyclic. This concludes that $|\text{MAIS}(\mathcal{D}')| = a + 1$.

We show that we can always construct $|\text{MAIS}(\mathcal{D}')| = a+1$ based on $|\text{MAIS}(\mathcal{D})| = a < s + t$. Therefore if there exists a $\mathcal{D}$ such that $|\text{MAIS}(\mathcal{D})| < s + t$, by the construction we have have a $\mathcal{D}'$ such that $|\text{MAIS}(\mathcal{D}')| = s + t - 1$. ∎

We are now ready to prove Proposition 6.

C. Proof of Proposition 6

Our proof for Proposition 6 is by contradiction. Specifically, we prove that, under the assumption that there exists $\mathcal{D}'$ such that $|\text{MAIS}(\mathcal{D}')| = s + t - 1$ (see Property MAIS-P2) and given a valid code, there must exist a user that can decode

$$
\begin{array}{c|c|c|c|c|c||c|c|c}
 & w_1 & w_2 & w_3 & \dots & w_{s+1} & w_{s+2} & \dots & w_{2s+1} \\
\hline
s+1 \text{ users} & 0 & 1 & 1 & \dots & 1 & \multicolumn{3}{c}{B_1} \\
 & 1 & 0 & 1 & \dots & 1 & \multicolumn{3}{c}{B_2} \\
 & \dots & \dots & \dots & \dots & \dots & \multicolumn{3}{c}{\dots} \\
 & 1 & 1 & 1 & \dots & 0 & \multicolumn{3}{c}{B_{s+1}}
\end{array}
$$

Side Information | Decoded Messages

Fig. 2: Side information sets and decoded messages for $s+1$ users for Proposition 6.Case2.

$s+t$ messages. This however contradicts Property MAIS-P1. Therefore $\mathcal{D}'$ does not exist, which implies that there must exists a user that can decode $s+t$ messages and $|\text{MAIS}(\mathcal{D})| = s+t$ for all $\mathcal{D}$. This proves that for the critical case the optimal number of transmission is $\ell^* = s+t$.

Specifically, the assumption that $|\text{MAIS}(D')| = s+t-1$ implies that one can find a set of $s+t-1$ users, denoted by V, who desire different messages and with a strict partial order on V given by: for distinct $i, j \in V$, if $i < j$ then $d_j \notin A_i$. Without loss of generality, let $[s+2 : 2s+t]$ be the set of the distinct $s+t-1$ desired messages by the users in V. By the definition of MAIS, there is a user in V such that its side information set satisfies $A \subset [s+1]$. This is the user that has no incoming edges in the induced acyclic subgraph of the MAIS. Thus, a user with side information including the messages in $[s+1]$ (these messages are not desired by the users in V) is able to decode all the messages in $[s+2 : 2s+t]$.

Consider the following $s+1$ users: for $i \in [s+1]$ user u_i has side information $A_i = [s+1]\backslash\{i\}$. The side information sets and decoded messages of these users are illustrated in Fig. 2 where columns are for messages and rows for users; a 0 (resp. 1) entry indicates the absence (resp. presence) of the corresponding message in the side information set of the user. We have one of two cases:

Proposition 6.Case1: Assume that for some $k \in [s+1]$ we have $B_k \cap [s+1] = [s+1]\backslash A_k$. Recall B_k is the set of messages that user u_k can decode and A_k its side information, this user will gain the knowledge of all messages $W_{[s+1]}$. It therefore can decode all the remaining messages $W_{[s+2:2s+t]}$. Eventually this user decodes $s+t$ messages, therefore $C_k = [2s+t]$.

Proposition 6.Case2: For every user $i \in [s+1]$, we have $B_i \subseteq [s+2 : 2s+t]$–as shown in Fig. 2, where the side information and decodable message sets are represented by the rows of the matrix. The left part of the matrix indicates the side information of the users, where $0, 1$ entries show the absence and existence of the corresponding messages in the side information. By assumption $B_i \subseteq [s+2 : 2s+t]$ contains the indices of the messages decoded by user u_i and property BlockCover-P3, we have $|\cap_{i\in P} C_i| \notin [s : s+t-1]$ for any $P \subseteq [s+1]$. Note that $|\cap_{i\in P} A_i| = s+1-|P|$ and $A_i \cap B_i = \varnothing$, thus we have $|\cap_{i\in P} B_i| \notin [|P|-1 : |P|+t-2], \forall P \subseteq [s+1]$.

In Proposition 6.Case2, all $B_i,\ i \in [s+1]$ are non-empty subsets of a ground set $[s+2 : 2s+t]$; by Lemma 4 in the Appendix, it is guaranteed that there is a P such that $|[s+2 : 2s+1] \cap (\cap_{i\in P} B_i)| = |P|-1$; therefore we have $|\cap_{i\in P} B_i| \in [|P|-1 : |P|+t-2]$ for some $P \subseteq [s+1]$, which contradicts what we just stated, thus this case in impossible.

Therefore only Proposition 6.Case1 is possible. This shows the existence of a user whose block cover is $[m] = [2s+t]$. This user can decode $s+t$ messages. But this contradicts the assumption that the MAIS bound is $|\text{MAIS}(\mathcal{D}')| = s+t-1$. Overall, this shows that for all possible choices of $\mathcal{D}$ one must have $|\text{MAIS}(\mathcal{D})| \geqslant s+t$, which implies $\ell^* \geqslant s+t$. This, with the achievability in Proposition 1, concludes the proof of Proposition 6.

D. Complete–S where $|S| = 1$

With Proposition 6, we can prove a more general case.

Proposition 7. *(The case $|S| = 1$.) For the complete–$\{s\}$ PICOD(t) with m messages, the optimal code length is $\ell^* = \min\{s+t, m-s\}$.*

Proof of Proposition 7: Proposition 6 solves the case where $S = \{s\}$ and $m = 2s+t$. Therefore, in the following we study the remaining two cases: $m < 2s+t$ and $m > 2s+t$.

Case $m < 2s+t$: Consider an integer $\alpha \leqslant s$ and split the $n = \binom{m}{s}$ users in the system into two categories: users u_i with $[\alpha] \subset A_i$, and the other users. The users in the first category do not decode any message in $[\alpha]$ (since they have all these messages in their side information set); these users together form a complete–$\{s-\alpha\}$ PICOD(t) with $m-\alpha$ messages. Since this complete–$\{s-\alpha\}$ PICOD(t) is a subset of the original complete–$\{s\}$ PICOD(t), its optimal number of transmissions is a lower bound on the number of transmissions in the original system. If we take $m-\alpha = 2(s-\alpha)+t \iff \alpha = 2s+t-m > 0$ then, by Proposition 6, the optimal number of transmissions for the complete–$\{s-\alpha\}$ PICOD(t) with $m-\alpha = 2(s-\alpha)+t$ messages is $(s-\alpha)+t = m-s$. Therefore the original complete–$\{s\}$ PICOD(t) requires at least $m-s$ transmissions, i.e., $\ell^* \geqslant m-s = \min\{m-s, s+t\}$.

Case $m > 2s+t$: The proof is by contradiction. Assume there exists a D' such that $|\text{MAIS}(D')| = s+t-1$ and, without loss of generality, that the maximum acyclic induced subgraph is formed by users with desired messages $[s+t-1]$. Specifically, we have users $u_i, i \in [s+t-1]$ such that $d_i = i$ and $d_j \notin A_i$ for any $j, i \in [s], j > i$ (by the definition of MAIS and its induced partial order).

Let U' index the users whose side information is a subset of $[s+t : m]$, i.e., $i \in U'$ if $A_i \subset [s+t : m]$. Apparently $1 \in U'$. We distinguish the following two cases.

Proposition 7.($m > 2s+t$).Case1: If there is a user $u_i \in U'$ with desired message $d_i \in [s+t : m]$, we have $d_j \notin A_t$ for all $j \in [s]$. Thus users $u_i, u_1, u_2, \dots, u_{s+t-1}$ form an acyclic induced subgraph of length $s+t$. This contradicts to the assumption that $|\text{MAIS}(D')| = s+t-1$.

Proposition 7.($m > 2s+t$).Case2: For all $i \in U'$ we have $d_i \in [s]$. By a similar reasoning as in proof of Proposition 6, we can show that there exists a user who can decode $s+t$ messages. This again contradicts the assumption that $|\text{MAIS}(D)| = s+t-1$.

By combining Proposition 7.($m > 2s+t$).Case1 and Proposition 7.($m > 2s+t$).Case2, we conclude that $|\text{MAIS}(D)| > s$. By Properties MAIS-P1 and MAIS-P2 we thus have $\ell^* \geqslant s+t = \min\{m-s, s+t\}$.

The achievability follows directly the schemes in Proposition 1. Since $|S| = 1$, no partition is needed. ■

VI. Complete–S PICOD(t) where S is consecutive: Proof of Theorem 2

With Proposition 7, we are ready to prove Theorem 2 in full generality. We consider the following three cases.

A. Case $s_{\max} \leqslant \lceil (m-t)/2 \rceil$: $\ell^* = s_{\max} + t$

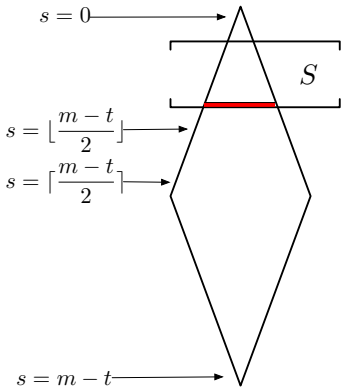

Fig. 3: $s_{\max} \leqslant \lceil (m-t)/2 \rceil - 1$.

Drop all the users except those with side information set of size $s_{\max}$, thereby obtaining a compete-$\{s_{\max}\}$ PICOD(t) with m messages. The layer representation of this case is shown in Fig. 3, where the red layer is the one left after dropping users. For this system the optimal number of transmissions is lower bound by $\min\{m - s_{\max}, s_{\max} + t\} = s_{\max} + t$ (because $s_{\max} \leqslant \lceil (m+t)/2 \rceil$ in this case), which is a lower bound on the number of transmissions in the original system. By Proposition 1, we have $\ell^* = s_{\max} + t$.

B. Case $s_{\min} \geqslant \lfloor (m-t)/2 \rfloor$: $\ell^* = m - s_{\min}$

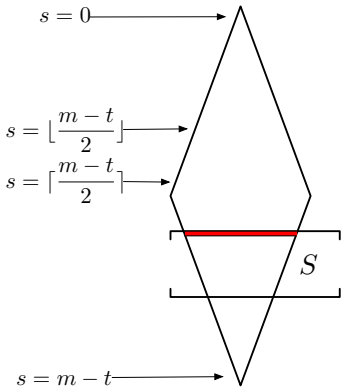

Fig. 4: $s_{\min} \geqslant \lfloor (m-t)/2 \rfloor$.

As for the case in Section VI-A, drop all the users except those with side information of size $s_{\min}$, thereby obtaining a compete-$\{s_{\min}\}$ PICOD(t) with m messages and the optimal number of transmissions is lower bounded $\min\{m - s_{\min}, s_{\min} + t\} = m - s_{\min}$ (because $s_{\min} \geqslant \lfloor (m-t)/2 \rfloor$ in this case). By Proposition 1, we have $\ell^* = m - s_{\min}$. The layer representation of this case is shown in Fig. 4, where red layer is the one left after dropping users.

C. Case $s_{\min} \leqslant \lfloor (m-t)/2 \rfloor - 1 \leqslant \lceil (m-t)/2 \rceil \leqslant s_{\max}$

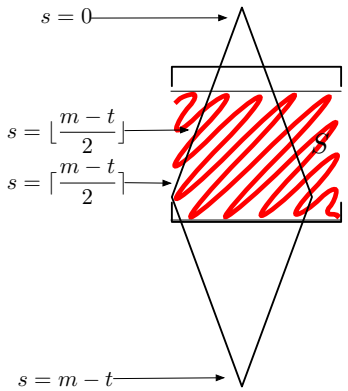

Fig. 5: $s_{\min} \leqslant \lfloor (m-t)/2 \rfloor - 1 \leqslant \lceil (m-t)/2 \rceil \leqslant s_{\max}$.

Define

$$\delta := \min\left\{ s_{\max} - \lceil \frac{m-t}{2} \rceil, \lfloor \frac{m-t}{2} \rfloor - s_{\min} \right\}, \tag{17}$$

$$m' := m + 2\delta + \lceil \frac{m-t}{2} \rceil - \lfloor \frac{m-t}{2} \rfloor, \tag{18}$$

$$S' := \left[\lfloor \frac{m-t}{2} \rfloor - \delta : \lceil \frac{m-t}{2} \rceil + \delta \right]. \tag{19}$$

Drop all users except those with side information of size $s \in S'$ for S' in (19), thereby obtaining a complete–S' PICOD(t) with m messages. The layer representation of this case is shown in Fig. 5, where red layers are the ones left after dropping users. Create dummy messages $W_{[m+1:m']}$, where dummy messages will not be desired by any user. To every user who was not dropped, with side information of size $s \in S'$, give every $(\lceil \frac{m-t}{2} \rceil + \delta - s)$-subset of $[m+1 : m']$ as extra side information (where δ is defined in (17) and m' in (18)); each such user generates $\binom{m'-m}{\lceil \frac{m-t}{2} \rceil + \delta - s}$ new users. All the users created by this procedure form a complete–$\{\lceil \frac{m-t}{2} \rceil + \delta\}$ PICOD(t) with m' messages, whose optimal number of transmissions is

$$\begin{aligned}
&\min\left\{ \lceil \frac{m-t}{2} \rceil + \delta + t, m' - (\lceil \frac{m-t}{2} \rceil + \delta) \right\} \\
&= \min\Bigg\{ \lceil \frac{m-t}{2} \rceil + \delta + t, \\
&\quad m + 2\delta + \lceil \frac{m-t}{2} \rceil - \lfloor \frac{m-t}{2} \rfloor - \lceil \frac{m-t}{2} \rceil - \delta \Bigg\} \\
&= \delta + t + \min\left\{ \lceil \frac{m-t}{2} \rceil, m - t - \lfloor \frac{m-t}{2} \rfloor \right\} \\
&= \delta + t + \lceil \frac{m-t}{2} \rceil \\
&= \min\left\{ s_{\max} - \lceil \frac{m-t}{2} \rceil, \lfloor \frac{m-t}{2} \rfloor - s_{\min} \right\} \\
&\quad + t + \lceil \frac{m-t}{2} \rceil \\
&= \min\left\{ s_{\max} + t, m - s_{\min} \right\} \\
&= \ell'.
\end{aligned}$$

Although the new system contains more users, any valid code for the original system works for the new one. Therefore the optimal code length ℓ' is a lower bound on the optimal code length for the original system. This lower bound can be attained by the scheme described in Proposition 1. This concludes Theorem 2.

VII. Some other complete–S PICOD(t)

The proofs in Section V-D start by dropping some users in the system. This shows that there exists *non-critical users* that do not affect the optimal code length. Therefore, by adding non-critical users, we can obtain a *non-consecutive* complete–S PICOD(t) where the proof used for Theorem 2 can still provide a tight converse.

A. Proof of Proposition 2

The converse depends only on the users with side information of size $s_{\max}$. The code that satisfies the complete–$\{s_{\max}\}$ PICOD(t), i.e., transmit $s_{\max}+t$ messages one at a time, also satisfies all the users with smaller size of side information.

B. Proof of Proposition 3

The converse depends only on the users with side information of size $s_{\min}$. The code that satisfies the complete–$\{s_{\min}\}$ PICOD(t), i.e., transmit $m-s_{\min}$ linearly independent linear combinations of all messages, also satisfies all the users with larger size of side information.

C. Proof of Proposition 4

The converse depends only on the users with side information of size in $\left[\lfloor\frac{m-t}{2}\rfloor-\delta:\lceil\frac{m-t}{2}\rceil+\delta\right]$. The code that satisfies the complete–$\left[\lfloor\frac{m-t}{2}\rfloor-\delta:\lceil\frac{m-t}{2}\rceil+\delta\right]$ PICOD(t) also satisfies all the users with larger size of side information set. That is, either transmit $s_{\max}+t$ messages one at a time, or $m-s_{\min}$ linearly independent linear combinations of all messages.

D. Proof of Proposition 5

Proposition 5 states that the achievable scheme in Proposition 1 is information theoretically optimal for the complete–S PICOD(t) with $m\leqslant 5$. The main idea behind these proofs follows the one in converse proof of Theorem 1: construct a decoding chain by providing proper messages to the user as genie, in a way that the user can mimic other users and decode the desired number of messages. Table I lists the optimal code length ℓ^* of all complete–S PICOD(t) instances that are not covered by Theorem 1 or Propositions 2, 3, 4.

Unfortunately, the converse proofs are based on a case-by-case reasoning, i.e., constructively find a user that can decode a certain number of messages. We could not straightforwardly extended these proof to the complete–S PICOD(t) for general m. Here we show proofs of two cases. The other cases can be proved using the similar methods.

TABLE I: Complete–S PICOD(t) that are not covered by Theorem 1 or Propositions 2, 3, 4.

$m=4$	$S=\{0,2\}$	$t=1,2$	$\ell^*=t+2$
	$S=\{1,3\}$	$t=1$	$\ell^*=3$
$m=5$	$S=\{0,3\}$	$t=1,2$	$\ell^*=t+2$
	$S=\{1,4\}$	$t=1$	$\ell^*=3$
	$S=\{1,3\}$	$t=1,2$	$\ell^*=4$
	$S=\{0,1,3\}$	$t=1,2$	$\ell^*=t+3$
	$S=\{1,3,4\}$	$t=1$	$\ell^*=4$
	$S=\{0,2,3\}$	$t=1,2$	$\ell^*=t+3$
	$S=\{0,2,4\}$	$t=1$	$\ell^*=4$
	$S=\{1,2,4\}$	$t=1$	$\ell^*=4$

a) Proposition 5.Case1: We show that for the complete–S PICOD(1) where $S=\{1,3\}$ and $m=5$, the optimal code has length $\ell^*=4$. We do so by proving the existence of a user with one message in its side information set who can decode the remaining 4 messages.

By Proposition 7, there exists a user, say u_1, with side information set of size 1, say $A_1=\{1\}$, who can decode 2 messages, say $B_1\supseteq\{2,3\}$. User u_1 thus can mimic user u_2 with side information $A_2=\{1,2,3\}$ and decode its desired message. Therefore user u_1 can decode at least 3 messages, $|B_1|\geqslant 3$.

Denote the last message that has not been decoded by user u_1 as w_5. Now, if w_5 is desired by some users, i.e., we have a user u_3 with $d_3=5$, user u_1 can mimic user u_3 and decode w_5 since $A_3\subset[4]$. Therefore user u_1 can decode 4 messages and $\ell^*\geqslant 4$.

Otherwise, w_5 is not desired by any users in the system. Since the message that is not desired by any users does not have any effect, by deleting it, the system becomes the complete–$[0:3]$ PICOD(1) with $m=4$. By Theorem 1 we have the user with $A=\{5\}$ can decode 4 messages and $\ell^*\geqslant 4$.

We apply the achievability for the complete–$\{1,2,3\}$ PICOD(1). This achievability works since $\{1,3\}\subset\{1,2,3\}$. By Theorem 2 we have $\ell^*\leqslant 4$. This proves the optimality of $\ell^*=4$ transmissions.

Note: the existence proof based on block cover, as used for Proposition 6, is also workable for Proposition 5 as well.

b) Proposition 5.Case2: We show that for the complete–S PICOD(1) problem where $S=\{0,2,4\}$ and $m=5$, the optimal code has length $\ell^*=4$. The following lemma, which is a refined version of Proposition 7, is used in the proof.

Lemma 2. *For a complete–$\{s\}$ PICOD(t) with m messages, let $A'\subset[m],|A'|\leqslant s$, $U_{A'}$ be the group of users who have A' in their side information, i.e., $u_i\in U_{A'}$ if and only if $A'\subseteq A_i$. For any A', there exists a user in $U_{A'}$ that can decode at least $\min\{m-s,s+t-|A'|\}$ messages. Note: Proposition 7 is the case $A'=\varnothing$.*

Proof of Lemma 2: The users in $U_{A'}$ alone can be seen as the users in a new complete–S' PICOD(t), where $S'=\{s-|A'|\}$, $m'=m-|A'|$. By Proposition 7 we have that there exists a user in this system that can decode $\min\{s'+t,m'-s'\}=\min\{s+t-|A'|,m-s\}$ messages. The above argument holds for all $A'\subset[m],|A'|\leqslant s$. ∎

Back to the proof of Proposition 5.Case2. We show that by giving one message as a genie, the user with no side information can decode the other 4 messages.

Since every user can decode one message, user u_1 with $A_1 = \varnothing$ can decode message w_{d_1}. By Lemma 2, we see that there exists a user $u_2 \in U_{\{d_1\}}$ that can decode 2 messages, where $U_{\{d_1\}}$ is the group of users who have side information sets of size 2 and w_{d_1} in their side information sets. Without loss of generality let $A_2 = \{d_1, 2\}$ and the two messages that u_2 can decode be w_3, w_4, $d_1 \notin \{2,3,4\}$. Therefore, giving message w_2 to user u_1 allows it to decode w_3, w_4. Also, there exists a user with side information $\{d_1, 2, 3, 4\}$ and decodes $w_{d_5} \notin \{d_1, 2, 3, 4\}$. So user u_1 can decode w_{d_5} as well. Overall, user u_1 can decode 4 messages with the proper genie w_2. The code length is therefore lower bounded by $\ell^* \geqslant 4$.

For the achievability, we split the users into two groups: $S_1 = \{0, 2\}$ where users have side information of size 0 or 2; $S_2 = \{4\}$ where users have side information of size 4. By Proposition 2 we can satisfy all users in S_1 with 3 transmission; by Proposition 7 we can satisfy all users in S_2 with one transmission. In total we use 4 transmissions to satisfy all users.

VIII. Proof of Theorem 3

In this section, we prove a tight converse bound on the optimal code length for PICOD(1) with circular-arc network topology hypergraph. We start by introducing some graph theory terminology.

A. Graph Preliminaries

Let $H = (V, \mathcal{E})$ denote a *hypergraph* with vertex set V and edge set $\mathcal{E}$, where an edge $E \in \mathcal{E}$ is a subset of V, i.e., $E \subseteq V$. The hypergraph is called *r-uniform* if all edges have cardinality r, i.e., $|E| = r$, $\forall E \in \mathcal{E}$. For $R \subseteq [|V|]$, the hypergraph is called *R-uniform* if all edges have cardinality of some $r \in R$, i.e., $|E| \in R$, $\forall E \in \mathcal{E}$. The hypergraph is called *complete r-uniform* if all edges with cardinality r exit, i.e., for all E such that $|E| = r, E \subseteq V$, we have $E \in \mathcal{E}$. The hypergraph is called complete *R-uniform* if all edges with cardinality $r \in R$ exist. The *dual* hypergraph $H^* = (V^*, \mathcal{E}^*)$ of H is a hypergraph where the vertices and edges are interchanged, i.e., $\mathcal{E}^* = V$, $V^* = \mathcal{E}$.

The degree of a vertex $v \in V$ is the number of its incident edges, i.e., $\delta(v) = |\{E : v \in E, E \in \mathcal{E}\}|$. The hypergraph is called *k-regular* if the degree of all vertices is k. A *factor* of H is a spanning edge induced subgraph of H, i.e., an edge induced subgraph of H with the same vertex set of V. A *k-factor* is a factor which is k-regular. A hypergraph H is called an *circular-arc hypergraph* if there exists an ordering of the vertices $v_1, v_2, \ldots, v_n$ such that if $v_i, v_j, i \leqslant j$, then the v_q for either all $i \leqslant q \leqslant j$, or all $q \leqslant i$ and $q \geqslant j$, are incident to an edge E.

For a PICOD(t), its network topology hypergraph is a hypergraph $H = (V, \mathcal{E})$ such that: i) $V = \{u_1, \ldots, u_n\}$, i.e., vertices represent the users; ii) $\mathcal{E} = \{E_1, \ldots, E_m\}$, i.e., edges represent the messages; iii) $u_i \in E_j$ if $w_j \notin A_i$, i.e., a vertex is incident to an edge if the user does not have the message in the side information. This definition of network topology hypergraph is a generalization of the network topology graph in [15].

Note that the network topology hypergraph is defined solely on user set U, message set W, and side information sets $\mathcal{A}$. For the IC, the network topology hypergraph does not uniquely define an instance of the problem, since it does not contain the information about desired message sets of the users. However, the network topology hypergraph uniquely defines a PICOD(t) due to the property that the PICOD(t) does not specify the desired messages for the users.

B. On the Optimality of a Single Transmission

We give the necessary and sufficient condition on the network topology hypergraph of a PICOD(1) problem for which one transmission is optimal. This result applies to all PICOD(1) instances, thus serves as a general converse bound for the PICOD(1).

Proposition 8. *A PICOD(1) with m messages has $\ell^* = 1$ if and only if its network topology hypergraph has a 1-factor. Otherwise $\ell^* \geqslant 2$.*

Proof of Proposition 8:

Achievability: The network topology hypergraph H has a 1-factor if it has an edge induced sub-hypergraph whose vertices are the same as the vertices of H and all have degree one. In other words, in this induced sub-hypergraph all vertices are adjacent to one and only one edge. Since H is the network topology hypergraph, its vertices represent users and edges represent messages. A vertex is adjacent to an edge if and only if the user does not have that message in its side information set. For the PICOD(1), that message can be a desired message by the incident users. Therefore, among all the messages corresponding to the edges in the 1-factor, every user has one and only one message that is not in its side information set. Transmitting the sum of all these messages satisfies all users. By this transmission scheme we achieve $\ell^* = 1$, which is clearly optimal.

Converse: We aim to show that if the network topology hypergraph does not have a 1-factor hypergraph, then we can construct a user that can decode two messages, thus two transmissions are needed. For any valid code, consider the sub-hypergraph induced by the edges corresponding to all the desired messages by all users, i.e., the edge induced sub-hypergraph of H where the edges correspond to the messages that are decoded by at least one user. This sub-hypergraph is always a factor, i.e., a spanning sub-hypergraph, since all users can decode at least one message in a PICOD(1). Assume no 1-factor exists in H, that is, there exists a vertex whose degree is at least 2 in the sub-hypergraph. In other words, for all choices of desired messages at the users, there exists a pair of users u_1 and u_2 with desired messages w_{d_1} and w_{d_2} such that $d_2 \notin A_1$. We therefore have $A_1 \subseteq [m] \backslash \{d_1, d_2\}$. Given any valid code, a user u' with $A' = [m] \backslash \{d_1, d_2\}$ can mimic user u_1 then user u_2, thus can decode w_{d_1}, w_{d_2}. By Lemma 1, we conclude that $\ell^* \geqslant 2$. ∎

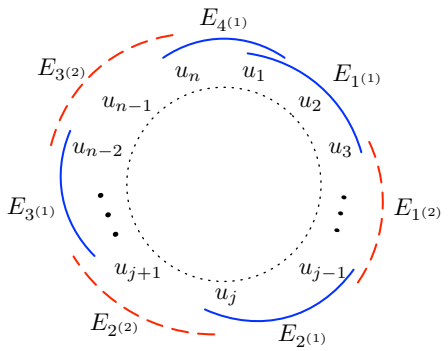

Fig. 6: Two transmissions scheme for circular-arc network topology hypergraph PICOD(t).

C. Proof of Theorem 3

We show a case where the converse proposed in Proposition 8 is tight by proposing an achievable scheme based on the properties of a circular-arc hypergraph. First, in Lemma 3 we show the following fact: if two edges, say E_i and E_j, are "close" in $\mathcal{E}$ with a nonzero gap between them, then there exists an edge in $\mathcal{E}$ that "covers" the whole gap between E_i and E_j. This fact will be used in Theorem 3 to design a two-transmission achievable scheme.

Lemma 3. *Assume a circular-arc hypergraph H without isolated vertices and where the vertices are in a cyclic order $\{v_1, v_2, \dots, v_n\}$. Assume there exist two edges $E_i = \{v_{i_1}, \dots, v_{i_p}\}$ and $E_j = \{v_{j_1}, \dots, v_{j_q}\}$ that satisfy the following two conditions: Condition1) $i_p + 1 < j_1$, and Condition2) every edge that contains any vertices in $\{v_{i_p+1}, \dots, v_{j_1-1}\}$ contains v_{i_p}. Then, there exists an edge E_k such that $\{v_{i_p+1}, \dots, v_{j_1-1}\} \subseteq E_k$.*

Proof of Lemma 3: Since H does not have any isolated vertices, there exists $E_k \in \mathcal{E}$ such that $v_{j_1-1} \in E_k$. By the Condition2 we have $v_{i_p} \in E_k$. By the property of circular-arc hypergraph (if v_{i_p} and v_{j_1-1} are contained in E_k, all the vertices between are contained in E_k as well) we have $\{v_{i_p+1}, \dots, v_{j_1-1}\} \subseteq E_k$. ∎

Proof of Theorem 3: We propose an achievable scheme that uses two transmissions to satisfy all users for all PICOD(1) instances with circular-arc network topology hypergraph. The scheme consists two steps.

a) Theorem 3.Step1: Given a PICOD(t) with network topology hypergraphas a circular-arc hypergraph, we notice that:

- No vertex is isolated.
- There might exists an edge that is as a proper subset of another edge.

We drop those edges that are proper subsets of the union of other edges, obtaining the edge set $\mathcal{E}$. In other words, $|E_i \backslash (\cup_{j \neq i} E_j)| > 0, \forall E_i, E_j \in \mathcal{E}$. The achievability scheme based on $\mathcal{E}$ will be valid for the original problem setting as well (since we are restricted to use less messages to satisfy all users). The edge induced subgraph by $\mathcal{E}$ has no isolated vertex as well.

Algorithm 1: Algorithm for finding $\mathcal{E}^{(1)}$ in Step1.

Data: User set: $V = \{v_1, \dots, v_n\}$, message set: $\mathcal{E}$.
Result: Message set: $\mathcal{E}^{(1)} = \{E_{1^{(1)}}, \dots E_{e^{(1)}}\}$.
Initialization: set $i = 1$, $\mathcal{E}^{(1)} = \varnothing$.
while $i \leqslant n$ **do**
  Seek an edge that starts at v_i, i.e., an edge that is $\{v_i, \dots\}$;
  if *Such an edge is found* **then**
    Let $\mathcal{E}^{(1)}$ include be the edge found;
    i becomes the index of the vertex right after the found edge, that is, the edge $\{\dots, v_{i-1}\}$;
  else
    $i = i + 1$;
  end
end

In Step1 we find a set of messages $\mathcal{E}^{(1)} \subseteq \mathcal{E}$ by using Algorithm 1. The blue solid arcs in Fig. 6 show an example of $\mathcal{E}^{(1)}$ found by Algorithm 1.

Denote the cardinality of $\mathcal{E}^{(1)}$ as $e := |\mathcal{E}^{(1)}|$. We claim that $\mathcal{E}^{(1)}$ has the following properties:

- $E_{i^{(1)}} \cap E_{j^{(1)}} = \varnothing$, for all $i, j \in [e], (i, j) \neq (1, e)$ and $(i, j) \neq (e, 1)$.
- For all $i, j \in [e-1]$, $E_{i^{(1)}} = \{v_{i_1^{(1)}}, \dots, v_{i_{e_i}^{(1)}}\}$, $E_{j^{(1)}} = \{v_{j_1^{(1)}}, \dots, v_{j_{e_j}^{(1)}}\}$, if $i_{e_i}^{(1)} + 1 < j_1^{(1)}$, we have an edge $E_{i^{(2)}} \in \mathcal{E}$ such that $\{v_{i_{e_i}^{(1)}}, \dots, v_{j_1^{(1)}}\}$.

The first property holds since the algorithm chooses adjacent edges in $\mathcal{E}^{(1)}$ that are disjoint and there is possibly nonempty intersection between $E_{1^{(1)}}$ and E_e. The second property holds by Lemma 3.

In the first transmission we send the sum of the messages in $\mathcal{E}^{(1)}$, i.e., $\sum_{i=1}^{e} w_i$. The users who are satisfied are in $\cup_{E_i \in \mathcal{E}^{(1)}} E_i \backslash (E_{1^{(1)}} \cap E_e)$. In the network topology hypergraph, these are the users that are "spanned" by these edges, excluding the users whose vertices are in $E_{1^{(1)}} \cap E_e$ where $E_{1^{(1)}} \cap E_e \neq \varnothing$. Therefore we are left with the users whose corresponding vertices are contained in $(U \backslash (\cup_{\mathcal{E}^{(1)}} E_{i^{(1)}})) \cup (E_{1^{(1)}} \cap E_e)$.

b) Theorem 3.Step2: The users who are not satisfied by the first transmission are the users whose side information sets contain either all the chosen messages in Theorem 3.Step1, or both $w_{1^{(1)}}$ and w_e. In other words, in the network topology hypergraph, they are the users who lie "in between" the edges, or in the intersection of the first and last edges, in $\mathcal{E}^{(1)}$ chosen in the previous step.

As we have shown in the second property of $\mathcal{E}^{(1)}$ in Theorem 3.Step1, for the unsatisfied users between $E_{i^{(1)}} \in \mathcal{E}^{(1)}$ and $E_{(i+1)^{(1)}} \in \mathcal{E}^{(1)}$, there exists an edge $E_{i(2)}$ that includes all those users. Therefore, we find a set of edges $\mathcal{E}^{(2)} = \{E_{1^{(2)}}, \dots, E_{(e-1)^{(2)}}\}$ such that $U \backslash (\cup_{\mathcal{E}^{(1)}} E_{i^{(1)}}) \subseteq \cup_{E_i \in \mathcal{E}^{(2)}} E_i$. In Fig. 6 they are the edges represented by the red dashed arcs. Note that all edges in $\mathcal{E}^{(2)}$ are pairwise disjoint, since if $E_{i^{(2)}} \cap E_{i+1^{(2)}} \neq \varnothing$ then we have $E_{i^{(1)}} \subseteq E_{i-1^{(2)}} \cup E_{i^{(2)}}$, i.e., $|E_{i^{(1)}} \backslash (\cup_{j \neq i^{(1)}} E_j)| = 0$. This is forbidden since we dropped the messages at the beginning of the Step1. Moreover, $(E_{1^{(1)}} \cap E_{e^1}) \cap E_{1^{(2)}} = \varnothing$ and $(E_{1^{(1)}} \cap E_{e^1}) \cap E_{(e-1)^{(2)}} = \varnothing$ by the same reasoning.

In the second transmission, we send the sum $\left(\sum_{j=1}^{e-1} w_{j^{(2)}}\right) + w_{1^{(1)}}$. The users that are not satisfied yet by the first transmission have all but one of the messages in $\{w_{1^{(2)}}, \dots, w_{k-1^{(2)}}, w_{1^{(1)}}\}$ in their side information sets. Therefore all the unsatisfied user after Theorem 3.Step1 can be satisfied by the second transmission. All the users are satisfied with two transmissions.

This, together with the converse in Proposition 8, concludes the proof of Theorem 3. ∎

IX. Conclusion and Future Works

In this paper we provided tight information theoretic converse bounds for some classes of PICOD(t) problems. The key idea for our converse is to show that for the PICOD(t) with a certain structure of the side information sets, regardless of the choice of desired message sets at the users, there exists a user that can decode a certain number of messages beside its t desired ones. We showed two methods to prove the existence of such a user: constructive proof and existence proof. The constructive proof works for the PICOD(t) with circular-arc network topology hypergraph, and for the complement-consecutive complete–S PICOD(t) with m messages where $S = [0 : m-t] \backslash [s_{\min} : s_{\max}], 0 < s_{\min} \leqslant s_{\max} < m-1$. The existence proof works for the consecutive complete–S PICOD(t) with m messages where $S = [s_{\min} : s_{\max}], 0 \leqslant s_{\min} \leqslant s_{\max} \leqslant m-1$.

The key idea for the existence proof was inspired by the similarity of the side information set structure of the consecutive complete–$\{s\}$ PICOD(t) to Steiner systems in combinatorial design. Combinatorial design studies the properties of a family of subsets, called blocks, that cover all s-element subsets of the same ground set; the results are usually established on the high symmetry of the structure of all s-element subsets. We introduced the idea of block cover as a tool for the converse proof, together with the classical MAIS for the IC problem. We solved first the critical complete–$\{s\}$ PICOD(t) with $m = 2s + t$ messages, where we showed that a block cover with maximum block size strictly less than $m = 2s + t$ does not exist. For the other considered cases, we showed that we can enhance the system to a critical one.

Open problems and future directions include:

- The main contribution of this work are methods to prove the existence of a user that can decode a certain number of messages: constructive and existence proofs. While the latter shows an advantage over the former on the complexity of the proof, it is based on the strong symmetric structure of the side information set of the users. Like combinatorial design, for the result to hold we need exactly all the s-element subsets of ground set $[m]$. Therefore, this method suits the complete–$\{s\}$ PICOD(t). For the other cases, we need some extra tools. We showed the proof for the consecutive complete–S PICOD(t) by a reduction to the critical case. However, it appears that not all the PICOD(t), even all complete–S PICOD(t), can be reduced in the same fashion without loss of optimality in terms of the code length. Therefore we still lack an efficient method to obtain a general optimal converse bound for the general PICOD(t). In Section VII-D we showed the optimality of the proposed achievability up to $m = 5$ for the complete–S PICOD(t). The converse is obtained by checking all the cases that are not covered by the Theorem 1 or Propositions 2, 3, 4. Therefore the method is not systematic and straightforwardly generalizable to general m. The information theoretical optimal code length for the general complete–S PICOD(t) with m messages is still open.
- We notice that in the complete–S PICOD(t) considered in this work, removing/adding some users does not change the optimal code length. In fact, in some cases (e.g., $S = [0 : m/2]$) roughly half of the users can be removed without affecting ℓ^*. These users can be considered as "non-critical", in contrast to other "critical" users who will change the optimal code length if removed/added. The PICOD(t) is called "critical" if all of its users are critical. We see the "critical" consecutive complete–S PICOD(t) are those with $m \geqslant s_{\min} + s_{\max} + t$. In other words, the ones with "small" size of side information/number of desired messages. In this case the optimal code length $s_{\max} + t$. For this setting, removing any single user reduces the optimal code length by 1. If $m < s_{\min} + s_{\max} + 1$, there are $\sum_{s=s_{\min}}^{s_{\max}} \binom{m}{s} - \binom{2m-2s_{\min}-1}{m-s_{\min}-1}$ non-critical users. It is worth mention that due to the symmetric structure of the complete–S PICOD(t) where $|S| = 1$, all users are essentially the same, i.e., all users are critical if any user is critical. The question about the critical users in the PICOD(t) is interesting because it shows the redundancy embedded in the system structure. The condition for a complete–S PICOD(t) to be critical, the number of its non-critical users, and in general, the condition to be critical for the general PICOD(t), are the topics of future works for the PICOD(t).
- In the PICOD formulation adopted in this work, the server broadcasts information to all users based on the knowledge all messages in the database. Another practically motivated scenario includes peer-to-peer/distributed models where users broadcast information based on their side information set. The converse bounds developed in this work are also converse bounds for peer-to-peer/distributed model with the same parameters. The open question is whether this "trivial" converse bound can be achieved. Surprisingly, it appears that for the consecutive and complement-consecutive complete–S PICOD(t) that we have solved, as long as the problem is "pliable,"

i.e., there are indeed multiple choices of desired messages that satisfy the users, then the tight results in this paper are tight for the peer-to-peer/distributed model. One of the open questions is to quantify the optimal code length is the non-pliable cases for the complete–S PICOD(t), where the problem reduces to a distributed index coding problem [11].

APPENDIX

Lemma 4. *For $s+1$ arbitrary subsets B_i from a ground set of size s, there exists a set $P \subseteq [s+1]$ such that $|\cap_{i\in P} B_i| = |P| - 1$.*

The proof of Lemma 4 is based on induction on s (the size of the ground set in this Lemma) and the following Lemma 5.

Lemma 5. *Let $B_1, B_2, \ldots, B_x$ are non-empty subsets of set $\{v_1, v_2, \ldots, v_y\}$, for some positive integers x, y. Let C_j be the collection of subsets that contain v_j, i.e., $v_j \in B_i$ if and only if $i \in C_j$. Let $c_j = |C_j|$. There always exists a pair (i,j) such that $\frac{c_j}{|B_i|} \geqslant \frac{x}{y}$ and $v_j \in B_i$.*

Proof of Lemma 5: Construct a $x \times y$ matrix W. $w_{ij} = 1/|B_i|$ if $v_j \in B_i$, otherwise $w_{ij} = 0$. Since $|B_i| \neq 0$ for all i, matrix W can be constructed. Note that the sum of each row is one. We have the summation of all elements in W is $\sum_{i\in[x],j\in[y]} w_{ij} = \sum_{i\in[x]}(\sum_{j\in[y]} w_{ij}) = x$, which is the number of rows. The summation of all elements in W can also be obtained by adding up the summation of the columns. Since there are y columns, there exists a column whose summation is no less than the average, i.e., there exists j such that

$$\sum_{k\in[x]} w_{kj} = \sum_{k:v_j\in B_k} \frac{1}{|B_k|} \geqslant \frac{x}{y}. \tag{20}$$

Let B_i be the smallest subset that contains v_j. We have

$$\sum_{k:v_j\in B_k} \frac{1}{|B_k|} \leqslant \sum_{k:v_j\in B_k} \frac{1}{|B_i|} = \frac{c_j}{|B_i|}. \tag{21}$$

Therefore, for the pair (i,j) we have $v_j \in B_i$ and

$$\frac{c_j}{|B_i|} \geqslant \frac{x}{y}. \tag{22}$$

■

Proof of Lemma 4: When $|B_i| = 0$ for some i, take $P = \{i\}$, we have $|\cap_{i\in P} B_i| = 0 = |P| - 1$. Lemma 4 is proven. Therefore we just need to consider the case where all B_i are non-empty.

For the initial case $s = 1$ the statement in Lemma 4 is true. It can be easily seen since $B_1 = B_2 = \{1\}$ (this is the only $s+1 = 2$ non empty subsets from a ground set of cardinality $s = 1$). Take $P = [2]$; we have $|\cap_{i\in[2]} B_i| = 1 = 2 - 1$.

Assume the statement in Lemma 4 is true for all $s \leqslant t-1$. We construct a P such that $|\cap_{i\in P} B_i| = |P| - 1$ for $s = t$. In Lemma 5, substitute x by $s+1$ and y by s, we have a pair (i,j) such that $j \in B_i$ and $\frac{c_j}{|B_i|} \geqslant \frac{s+1}{s}$, where $c_j = |C_j|$ and $C_j \subseteq [s+1]$ is the collection of subsets that contain j. By reordering the labels, without loss of generality, let $i = 1$ and $B_i = B_1 = [j]$. Since $\frac{c_j}{|B_1|} \geqslant \frac{s+1}{s} > 1$, we have $c_j > j$, $|C_j \backslash \{1\}| > j - 1$. Consider $B'_{i'} := B_{i'} \cap [j-1]$, $i' \in C_j \backslash \{1\}$ where $B'_{i'}$ are subsets of $[j-1]$. Since $j - 1 < s$, by the inductive hypothesis there exists P' such that $|\cap_{i'\in P'} B'_{i'}| = |P'| - 1$. Let $P = P' \cup \{1\}$. Note that $j \in B_q$ for all $q \in P$ and $k \notin \cap_{q\in P} B_q$ for all $k \in [j+1 : s]$. We have $\cap_{q\in P} B_q = \cap_{j'P'} \cup \{j\}$. Then $|\cap_{q\in P} B_q| = |P'| - 1 + 1 = |P| - 1$ as $|P| = |P'| + 1$.

Therefore we can always find a P such that $|\cap_{i\in P} B_i| = |P| - 1$ for all positive integer s. ■

REFERENCES

[1] F. Arbabjolfaei and Y.-H. Kim, *Fundamentals of Index Coding*. Now Publishers, 2018.

[2] A. S. Avestimehr, A. Sezgin, and D. N. Tse, "Approximate capacity of the two-way relay channel: A deterministic approach," *46th Annual Allerton Conference on Communication, Control, and Computing*, 2008.

[3] Z. Bar-Yossef, Y. Birk, T. S. Jayram, and T. Kol, "Index coding with side information," *IEEE Transactions on Information Theory*, vol. 57, no. 3, pp. 1479–1494, Mar 2011.

[4] Y. Birk and T. Kol, "Informed-source coding-on-demand (ISCOD) over broadcast channels," *Proc. IEEE 17th INFOCOM*, pp. 1257–1264, 1998.

[5] S. Brahma and C. Fragouli, "Pliable index coding," *IEEE Transactions on Information Theory*, vol. 61, no. 11, pp. 6192–6203, Nov 2015.

[6] R. M. Karp, "Reducibility among combinatorial problems," *in Complexity of Computer Computations*, pp. 85–103, 1972.

[7] T. Liu and D. Tuninetti, "Pliable index coding: Novel lower bound on the fraction of satisfied clients with a single transmission and its application," *Information Theory Workshop (ITW)*, 2016.

[8] ——, "Information theoretic converse proofs for some picod problems," *Information Theory Workshop (ITW)*, 2017.

[9] E. Lubetzky and U. Stav, "Nonlinear index coding outperforming the linear optimum," *IEEE Transactions on Information Theory*, vol. 55, no. 8, pp. 3544–3551, August 2009.

[10] S. E. Rouayheb, A. Sprintson, and C. Georghiades, "On the index coding problem and its relation to netowrk coding and matroid theory," *IEEE Transactions on Information Theory*, vol. 56, no. 7, pp. 3187–3195, July 2010.

[11] P. Sadeghi, F. Arbabjolfaei, and Y.-H. Kim, "Distributed index coding," *Proc. Int. Symp. Inf. Theory*, 2016.

[12] L. Song and C. Fragouli, "A polynomial-time algorithm for pliable index coding," *IEEE Transactions on Information Theory*, vol. 64, no. 2, pp. 979 – 999, Feb 2018.

[13] H. Sun and S. A. Jafar, "Index coding capacity: How far can one go with only shannon inequalities?" *IEEE Transactions on Information Theory*, vol. 61, no. 6, pp. 3041–3055, June 2015.

[14] J. H. van Lint, "On the number of blocks in a generalized steiner system," *Journal of Combinatorial Theory*, vol. A, no. 80, pp. 353 – 355, 1997.

[15] X. Yi, H. Sun, S. A. Jafar, and D. Gesbert, "TDMA is optimal for all-unicast dof region of TIM if and only if topology is chordal bipartite," *IEEE Transactions on Information Theory*, vol. 64, no. 3, pp. 2065 – 2076, Mar 2018.