A Note on Parity Check Matrices of Private Information Retrieval Codes

1st Koki Kazama
Shonan Institute of Technology
Email: kazama@info.shonan-it.ac.jp

2nd Takahiro Yoshida
Nihon University
yoshida.takahiro@nihon-u.ac.jp

***Abstract*—A private information retrieval (PIR) is an information retrieval scheme that allows a user to retrieve messages from information databases while keeping secret which one the user wants to retrieve. Sun et al. formulated the download rate and showed that there is an upper limit to it (capacity). Previous construction methods of a capacity-achieving linear PIR (CALPIR) are ad hoc. We propose conditions for a CALPIR using the parity check matrix.**

***Index Terms*—private information retrieval**

The full version of this paper is in [1].

I. Introduction

A private information retrieval (PIR) is a scheme that allows a user to retrieve messages from an information database (DB) while keeping secret which one the user wants to retrieve (user's privacy). One scheme is for the user to download all information from all DBs. However, the communication efficiency of this scheme is low. We construct a PIR with high communication efficiency. Chor et al. first studied a PIR [2]. Since Sun et al. studied a PIR [3] using information theory and coding theory, numerous studies have followed Sun's study.

The scheme of a PIR [3] and this study is as follows: Multiple databases store multiple messages. First, the user generates queries and sends them to the DBs. Then, each DB generates responses based on the query and all messages and sends a response to the user. Lastly, the user reconstructs the message that the user wants to retrieve. Each DB independently tries to guess which message the user wants when receiving a query, while each DB cannot obtain any information about it.

Sun et al. [3] formulated the user's privacy and a download rate, which is the communication efficiency when a user receives responses from all DBs, using information theory and showed that there exists an upper limit (PIR capacity) to the download rate when we fix the numbers of the DBs and the messages. The representative construction methods of a PIR that achieves the capacity and whose response-generating function (encoder) and reconstruction function (decoder) are linear (capacity-achieving linear PIR, CALPIR) are the study by Sun et al. [3] and that by Tian et al. [4]. However, both were ad hoc because describing the converse part of the proof of the PIR capacity using entropy did not clarify the conditions of the encoder and the decoder in the form of matrices.

We show the conditions of a CALPIR by its parity check matrix to propose a systematic construction method of a large subclass of CALPIRs. The reason why we want such construction methods is that there must be a number of construction methods for PIRs, and such construction methods may immediately suggest a variety of PIRs that have advantages under different evaluation criteria. Furthermore, using parity check matrices make conditions for generator matrices of CALPIRs expilicit since the parity check matrices decompose complex conditions that we want CALPIRs to satisfy into several simple conditions and each block row of the parity check matrices directly represents those conditions. Although several previous studies [5] [6] considered parity check matrices of PIRs, the policies differs from that of this study.

II. Reformulation of Previous Studies: The Capacity of PIRs

This section reformulates the previous studies of PIRs generally. Specifically, an encoder, a decoder and a decoder matrix are original to this study.

A. Problem Setting of PIRs

We explain the problem setting of a PIR concerning [3] [4].

First, we define notations. $\mathbb{N} \coloneqq \{1, 2, 3, \dots\}$. $[m, n] \coloneqq \{\, m, m+1, \dots, n \,\}$ and $[n] \coloneqq [1, n]$ for any $m, n \in \mathbb{N}$. All vectors are row vectors except specifically noted. $\boldsymbol{E}^\top$ denotes the transpose of a matrix $\boldsymbol{E}$. $\mathbb{F}_p$ is a finite field with p elements. $\mathbb{F}_p^{n\times m}$ denotes the set of all $n \times m$ matrices over $\mathbb{F}_p$, and $\mathbb{F}_p^n \coloneqq \mathbb{F}_p^{1\times n}$. For any $\boldsymbol{a} = (a_1, \dots, a_n)$ and $A \subset [n]$, $i \in [n]$, we define $\boldsymbol{a}_A \coloneqq (a_i)_{i\in A}$ and $\boldsymbol{a}_i \coloneqq a_i = \boldsymbol{a}_{\{i\}}$. $\mathrm{Uni}(A)$ denotes the uniform distribution on a finite set A. $\boldsymbol{I}_n$ denotes the $n \times n$ identity matrix. The base of the logarithm in Shannon entropy H is always p.

Next, we define an information retrieval (IR) and a PIR.

Definition 2.1 ((replicated) PIR [3] [4] [7]): Let $L_{\mathrm{w}}, L_{\mathrm{a}}, M, N, S$ be positive integers with $M, N \geq 2$. Let p be a prime power. Let $\boldsymbol{W} = (W_1, \dots, W_M)$ be a random variable vectors such that $W_1 = (W_{1,1}, \dots, W_{1,L_{\mathrm{w}}}), \dots, W_M = (W_{M,1}, \dots, W_{M,L_{\mathrm{w}}}) \overset{\text{i.i.d}}{\sim} \mathrm{Uni}\left(\mathbb{F}_p^{L_{\mathrm{w}}}\right)$. Let $\boldsymbol{S} (\sim \mathrm{Uni}([S]))$ be a random variable independent of $\boldsymbol{W}$. Let $\phi_n \colon \mathbb{N} \times \mathbb{F}_p^{L_{\mathrm{w}}M} \to \mathbb{F}_p^{L_{\mathrm{a}}}$ for $n \in [N]$, $\chi \colon [M] \times [S] \to \mathbb{N}^N$ and $\psi \colon [M] \times \mathcal{Q}^N \times \mathbb{F}_p^{L_{\mathrm{a}}N} \to \mathbb{F}_p^{L_{\mathrm{w}}}$ be functions. For any $\boldsymbol{q} = (q_1, \dots, q_N)$, we define $R(\boldsymbol{q}) \coloneqq \{\, (n, q_n) \mid n \in [N] \,\}$.

There are N DBs and each DB stores a *local encoder* ϕ, a *local decoder* ψ, and a *message vector* $\boldsymbol{W}$. M is the *message* *length* and L_{w} is the *message symbol length*. The user retrieves W_m by 3 phases, where $m(\in [M])$ is a *desired index*.

$\langle$*Query Phase*$\rangle$ The user generates a *query vector* $(Q_1^{(m)}, \dots, Q_N^{(m)}) := \chi(m, \boldsymbol{S})$ and send a *query* $Q_n^{(m)}$ to each DB $n \in [N]$.

$\langle$*Answer Phase*$\rangle$ Each DBn sends a *responce* $A_{\left(n, Q_n^{(m)}\right)} := \phi_n \left(Q_n^{(m)}, \boldsymbol{W} \right)$ to the user.

$\langle$*Decoding Phase*$\rangle$ The user decodes W_m from m, S and the *responce vector*$\boldsymbol{A}_{R\left(\boldsymbol{\mathcal{Q}}^{(m)}\right)} := \left(\phi_n \left(Q_n^{(m)}, \boldsymbol{W} \right) \right)_{n \in [N]}$ without errors. i.e.

$$\psi \left(m, \boldsymbol{S}, \left(\phi_n \left(\chi_n(m, \boldsymbol{S}), \boldsymbol{W} \right) \right)_{n \in [N]} \right) = W_m, \tag{1}$$

where $\chi_n(m, \boldsymbol{S}) = Q_n^{(m)}$ is the n th entry of for $n \in [N]$. We define an *information retrieval (IR)* with the parameters $(M, N, L_{\mathrm{w}}, L_{\mathrm{a}}, p)$ as this protocol. Hereafter we omit "with the parameters $(M, N, L_{\mathrm{w}}, L_{\mathrm{a}}, p)$." We define an *IR code* as a tuple (ϕ, ψ, χ, S) in the IR. Moreover, we define a *private information retrieval (PIR)* as an IR satisfying the *(user's) privacy constraint* Eq.(2).

$$\begin{aligned} &\forall m \in [M], \forall m' \in [M], \forall n \in [N], \forall q \in \mathbb{N}, \\ &\Pr \left(Q_n^{(m)} = q \right) = \Pr \left(Q_n^{(m')} = q \right). \end{aligned} \tag{2}$$

For simplicity, we assume Assumption2.1.

Assumption 2.1: A function $\chi(m, \cdot)$ is injective for any $m \in [M]$. $\mathcal{Q}_n := \left\{ q \in \mathbb{N} \,\middle|\, \Pr \left(Q_n^{(m)} = q \right) > 0 \right\} = \{ \chi_n(m, \boldsymbol{s}) \mid \boldsymbol{s} \in [S] \}$ does not depend on m for any IR.

The first sentence in Assumption 2.1 shows that the last equaltion in Eq.(2) is equivalent to

$$\left| \left\{ (q_1, \dots, q_N) \in \boldsymbol{\mathcal{Q}}^{(m)} \,\middle|\, q_n = q \right\} \right| = \left| \left\{ (q_1, \dots, q_N) \in \boldsymbol{\mathcal{Q}}^{(m')} \,\middle|\, q_n = q \right\} \right| \tag{3}$$

where $\boldsymbol{\mathcal{Q}}^{(m)} := \{ \chi(m, \boldsymbol{s}) \mid \boldsymbol{s} \in [S] \}$ (the m-th *query vector set*). This is a slight modification of Condition 11 in [8].

The second in Assumption 2.1 is valid if an IR is a PIR.

Definition 2.2 (codeword): We define a *codeword* of the IR as a random variable vector $(\phi_n(q, \boldsymbol{W}))_{n \in [N], q \in \mathcal{Q}_n}$.

The idea of considering IRs as codes appeared in [7] [9] although codeword indices are different with this study.

Definition 2.3 (encoder, decoder): For an IR, we define functions $\Phi \colon \mathbb{F}_p^{L_{\mathrm{w}} M} \to \mathbb{F}_p^{L_{\mathrm{a}} \sum_{n \in [N]} |\mathcal{Q}_n|}$ (*encoder*) and $\Psi \colon [M] \times [S] \times \mathbb{F}_p^{L_{\mathrm{a}} \sum_{n \in [N]} |\mathcal{Q}_n|} \to \mathbb{F}_p^{L_{\mathrm{w}} MS}$ (*decoder*) as

$$\Phi(\boldsymbol{w}) := (\phi_n(q, \boldsymbol{w}))_{n \in [N], q \in \mathcal{Q}_n}, \tag{4}$$

$$\Psi(\boldsymbol{a}) := \left(\psi \left(m, \boldsymbol{s}, \boldsymbol{a}_{R(\chi(m, \boldsymbol{s}))} \right) \right)_{m \in [M], \boldsymbol{s} \in [S]} \tag{5}$$

for any $\boldsymbol{w} \in \mathbb{F}_p^{L_{\mathrm{w}} M}$ and $\boldsymbol{a} \in \mathbb{F}_p^{L_{\mathrm{a}} \sum_{n \in [N]} |\mathcal{Q}_n|}$.
Then, from Eq.(1), it holds that for any $\boldsymbol{w} \in \mathbb{F}_p^{L_{\mathrm{w}} M}$,

$$\Psi(\Phi(\boldsymbol{w})) = (\underbrace{w_1, \dots, w_1}_{S}, \dots, \underbrace{w_M, \dots, w_M}_{S}). \tag{6}$$

B. Linear PIR

We redefine a linear PIR (LPIR) concerning [4], [8].

Definition 2.4 (a linear PIR [4] [8]): We define a *linear IR (LIR)* as an IR such that there exists matrices $\boldsymbol{G}_{(n, \chi_n(m, \boldsymbol{s}))} \in \mathbb{F}_p^{ML_{\mathrm{w}} \times L_{\mathrm{a}}}$ (*local generator matrix*) and $\boldsymbol{D}_{\boldsymbol{s}, \cdot}^{(m)} \in \mathbb{F}_p^{L_{\mathrm{w}} \times NL_{\mathrm{a}}}$ (*local decoder matrix*) such that $\phi_n(\chi_n(m, \boldsymbol{s}), \boldsymbol{w}) = \boldsymbol{w} \boldsymbol{G}_{(n, \chi_n(m, \boldsymbol{s}))}$ and $\psi(m, \boldsymbol{s}, \boldsymbol{a}') = \boldsymbol{a}' \left(\boldsymbol{D}_{\boldsymbol{s}, \cdot}^{(m)} \right)^{\top}$ for any $n \in [N], \boldsymbol{s} \in [S], \boldsymbol{w} \in \mathbb{F}_p^{L_{\mathrm{w}} M}, \boldsymbol{a}' \in \mathbb{F}_p^{L_{\mathrm{a}} N}$. A PIR that is an LIR is referred to as a *linear PIR* (LPIR).

Definition 2.5 (generator matrix [4], decoder matrix [8]): For an LIR, there exist a matrix $\boldsymbol{G} \in \mathbb{F}_p^{ML_{\mathrm{w}} \times L_{\mathrm{a}} \sum_{n \in [N]} |\mathcal{Q}_n|}$ (*generator matrix*) and a matrix $\tilde{\boldsymbol{D}} \in \mathbb{F}_p^{L_{\mathrm{w}} MS \times L_{\mathrm{a}} \sum_{n \in [N]} |\mathcal{Q}_n|}$ (*decoder matrix*) such that for any $\boldsymbol{w} \in \mathbb{F}_p^{ML_{\mathrm{w}}}$ and $\boldsymbol{a} \in \mathbb{F}_p^{L_{\mathrm{a}} \sum_{n \in [N]} |\mathcal{Q}_n|}$, it holds that $\Phi(\boldsymbol{w}) = \boldsymbol{w} \boldsymbol{G}$ and $\Psi(\boldsymbol{a}) = \boldsymbol{a} \tilde{\boldsymbol{D}}^{\top}$.

We construct the generator matrix from the local generator matrices and the decoder matrix from the local decoding matrices. $\boldsymbol{G} := \begin{pmatrix} \boldsymbol{G}_{(1,1)}^{(1)} & \cdots & \boldsymbol{G}_{(N, |\mathcal{Q}_N|)}^{(1)} \\ \vdots & \ddots & \vdots \\ \boldsymbol{G}_{(1,1)}^{(M)} & \cdots & \boldsymbol{G}_{(N, |\mathcal{Q}_N|)}^{(M)} \end{pmatrix}$, where $\boldsymbol{G}_{(n,q)}^{(m)} \in \mathbb{F}_p^{L_{\mathrm{w}} \times L_{\mathrm{a}}}$ for any $m \in [M], n \in [N], q \in \mathcal{Q}_n$. Each block column is $\boldsymbol{G}_{(n,q)} := \begin{pmatrix} \boldsymbol{G}_{(n,q)}^{(1)} \\ \vdots \\ \boldsymbol{G}_{(n,q)}^{(M)} \end{pmatrix}$ and each block row is $\boldsymbol{G}^{(m)} := \begin{pmatrix} \boldsymbol{G}_{(1,1)}^{(m)} & \cdots & \boldsymbol{G}_{(N, |\mathcal{Q}_N|)}^{(m)} \end{pmatrix}$. We define $\boldsymbol{G}_{R(\boldsymbol{q})} := \left(\boldsymbol{G}_{(n, q_n)} \right)_{n \in [N]}$ for any $\boldsymbol{q} = (q_1, \dots, q_N) \in \boldsymbol{\mathcal{Q}}^{(m)}$. We devide $\boldsymbol{D}_{\boldsymbol{s}, \cdot}^{(m)} = \begin{pmatrix} \boldsymbol{D}_{\boldsymbol{s}, 1}^{(m)} & \dots & \boldsymbol{D}_{\boldsymbol{s}, N}^{(m)} \end{pmatrix}$ so that $\boldsymbol{D}_{\boldsymbol{s}, n}^{(m)} \in \mathbb{F}_p^{L_{\mathrm{w}} \times L_{\mathrm{a}}}$. Then $\boldsymbol{w} \boldsymbol{G}_{R(\chi(m, \boldsymbol{s}))} \left(\boldsymbol{D}_{\boldsymbol{s}, \cdot}^{(m)} \right)^{\top} = w_m$ for any $\boldsymbol{w}, m, \boldsymbol{s}$. Moreover, we define

$$\tilde{\boldsymbol{D}} = \begin{pmatrix} \tilde{\boldsymbol{D}}^{(1)} \\ \vdots \\ \tilde{\boldsymbol{D}}^{(M)} \end{pmatrix}, \ \tilde{\boldsymbol{D}}^{(m)} = \begin{pmatrix} \tilde{\boldsymbol{D}}_{1, \cdot}^{(m)} \\ \vdots \\ \tilde{\boldsymbol{D}}_{S, \cdot}^{(m)} \end{pmatrix} \tag{7}$$

$$\tilde{\boldsymbol{D}}_{\boldsymbol{s}, \cdot}^{(m)} = \begin{pmatrix} \tilde{\boldsymbol{D}}_{\boldsymbol{s}, (1,1)}^{(m)} & \dots & \tilde{\boldsymbol{D}}_{\boldsymbol{s}, (N, |\mathcal{Q}_N|)}^{(m)} \end{pmatrix} \tag{8}$$

$$\mathbb{F}_p^{L_{\mathrm{w}} \times L_{\mathrm{a}}} \ni \tilde{\boldsymbol{D}}_{\boldsymbol{s}, (n,q)}^{(m)} := \begin{cases} \boldsymbol{D}_{\boldsymbol{s}, n}^{(m)} & \text{if } q = \chi_n(m, \boldsymbol{s}) \\ \boldsymbol{0}_{L_{\mathrm{w}} \times L_{\mathrm{a}}} & \text{otherwise.} \end{cases} \tag{9}$$

Lemma 2.1 (reconstruction of an LIR): We define $\boldsymbol{J} \in \mathbb{F}_p^{L_{\mathrm{w}} MS \times ML_{\mathrm{w}}}$ as the transpose of a matrix in Eq.(10), where $\boldsymbol{0} \in \mathbb{F}_p^{L_{\mathrm{w}} \times L_{\mathrm{w}}}$.

$$\begin{pmatrix} \boldsymbol{I}_{L_{\mathrm{w}}} & \dots & \boldsymbol{I}_{L_{\mathrm{w}}} & \dots & \boldsymbol{0} & \dots & \boldsymbol{0} \\ \boldsymbol{0} & \dots & \boldsymbol{0} & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \boldsymbol{0} & \dots & \boldsymbol{0} \\ \boldsymbol{0} & \dots & \boldsymbol{0} & \dots & \boldsymbol{I}_{L_{\mathrm{w}}} & \dots & \boldsymbol{I}_{L_{\mathrm{w}}} \end{pmatrix} \tag{10}$$

$$\underbrace{\qquad\qquad\qquad}_{L_{\mathrm{w}} S} \qquad\qquad \underbrace{\qquad\qquad\qquad}_{L_{\mathrm{w}} S}$$

Then for an LIR, it holds that

$$\tilde{\boldsymbol{G}} \begin{pmatrix} -\boldsymbol{J} & \tilde{\boldsymbol{D}} \end{pmatrix}^\top = \boldsymbol{0}_{ML_{\mathrm{w}} \times L_{\mathrm{w}} MS}, \tag{11}$$

where $\tilde{\boldsymbol{G}} := (\boldsymbol{I}_{ML_{\mathrm{w}}}, \boldsymbol{G})$ and $\boldsymbol{G}$ is a generator matrix.

This lemma is a modification of Propositon 9 in [8].

Proof : Eq. (6) is equivalent to $\boldsymbol{G}\tilde{\boldsymbol{D}}^\top = \boldsymbol{J}^\top$. □

C. PIR Capacity

The communication efficiency of an IR (not necessarily an LPIR) is the (download) rate. We do not discuss uploads.

Definition 2.6 (rate [3]): We define the rate of an IR as

$$\mathrm{rate}_{\mathrm{I}} := \min_{m \in [M]} \frac{L_{\mathrm{w}}}{\mathrm{H}\left(\boldsymbol{A}^{(m)} | \boldsymbol{Q}^{(m)}\right)}. \tag{12}$$

The rate is also equal to $\min_{m \in [M]} \frac{L_{\mathrm{w}} S}{\sum_{\boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}} \mathrm{H}\left(\Phi(\boldsymbol{W})_{R(\boldsymbol{q})}\right)}$.

Proposition 2.1 (Capacity [3]): We fix a finite field $\mathbb{F}_p$ and integers $M, N \geq 2$. For any PIR, it holds that

$$\mathrm{rate}_{\mathrm{I}} \leq \mathrm{C} := \left(1 - \frac{1}{N}\right) / \left(1 - \frac{1}{N^M}\right). \tag{13}$$

We define the *PIR capacity* as the R.H.S. of this iniquality. There exist $L_{\mathrm{w}}, L_{\mathrm{a}} \in \mathbb{N}$ and an LPIR that achieves equality in the iniquality (13) (*Capacity-Achieving LPIR, CALPIR*).

The proof is in Appendix A in [1].

Hereafter we fix the values of M, N and p. Condition 10 in [8] showed several capacity-achieving conditions of LPIRs.

III. This Study: Conditions of A Capacity-Achieving Linear PIR

We explain the conditions under which LPIR is always correctly decodable (Eq. (11)) and achieves the PIR capacity using its extended parity check matrix (EPCM). We already showed the user's privacy condition in Eq.(3). We omit numerical examples of previous studies due to the lack of space.

Definition 3.1 (extended parity check matrix of an LIR): We fix a LIR and a generator matrix $\boldsymbol{G}$ of this LIR. We define an *extended generator matrix* (EGM) of $\boldsymbol{G}$ as $\tilde{\boldsymbol{G}} := (\boldsymbol{I}_{ML_{\mathrm{w}}}, \boldsymbol{G})$. We define an *extended parity check matrix* (EPCM) of $\boldsymbol{G}$ as $\tilde{\boldsymbol{H}} \in \mathbb{F}_p^{h \times (L_{\mathrm{a}} \sum_{n \in [N]} |\mathcal{Q}_n| + ML_{\mathrm{w}})}$ such that $\mathrm{rank}\tilde{\boldsymbol{H}} = L_{\mathrm{a}} \sum_{n \in [N]} |\mathcal{Q}_n|$ and $\tilde{\boldsymbol{G}}\tilde{\boldsymbol{H}}^\top = \boldsymbol{0}_{ML_{\mathrm{w}} \times h}$, where $h \geq L_{\mathrm{a}} \sum_{n \in [N]} |\mathcal{Q}_n|$.

A. Conditions of Decoding

Definition 3.2: When two matrices $\boldsymbol{A}, \boldsymbol{B}$ are transferred to each other by a row basis transformation, the matrices $\boldsymbol{A}, \boldsymbol{B}$ are said to be equivalent and this is denoted by $\boldsymbol{A} \sim \boldsymbol{B}$.

Lemma 3.1: If a matrix $\tilde{\boldsymbol{H}}$ satisfies Eq.(14), $\tilde{\boldsymbol{H}}$ is an EPCM of $\boldsymbol{G}$.

$$\tilde{\boldsymbol{H}} \sim \begin{pmatrix} \boldsymbol{G}^\top & -\boldsymbol{I}_{L_{\mathrm{a}} \sum_{n \in [N]} |\mathcal{Q}_n|} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}. \tag{14}$$

Hereafter we focus on a construction method of $\tilde{\boldsymbol{H}}$ satisfying Eq.(14) and the decoder matrix $\tilde{\boldsymbol{D}}$.

Lemma 3.2: For a decoder matrix of an LIR $\tilde{\boldsymbol{D}}$, if a matrix $\tilde{\boldsymbol{H}}$ satisfies Condition 3.1, $\tilde{\boldsymbol{H}}$ is an EPCM of $\boldsymbol{G}$ in this condition.

Condition 3.1: There exist $h' \geq 0$, $\tilde{\boldsymbol{H}}' \in \mathbb{F}_p^{h' \times (ML_{\mathrm{w}} + L_{\mathrm{a}} \sum_{n \in [N]} |\mathcal{Q}_n|)}$ and $\boldsymbol{G} \in \mathbb{F}_p^{ML_{\mathrm{w}} \times L_{\mathrm{a}} \sum_n |\mathcal{Q}_n|}$ satisfying $\tilde{\boldsymbol{H}} = \begin{pmatrix} -\boldsymbol{J} & \tilde{\boldsymbol{D}} \\ \multicolumn{2}{c}{\tilde{\boldsymbol{H}}'} \end{pmatrix}$ and Eq.(14).

Theorem 3.1: If there exists a function χ and a matrix $\tilde{\boldsymbol{H}}$ satisfying Condition 3.2, we can construct an LIR whose generator matrix is $\boldsymbol{G}$ in Condition 3.1. Moreover, if they satisfy Eq.(3), the LIR is an LPIR.

Condition 3.2: χ and $\tilde{\boldsymbol{H}}$ satisfy Condition 3.1, where $\tilde{\boldsymbol{D}}$ is in the form of Eq.(7) and $\tilde{\boldsymbol{D}}^{(m)}_{\boldsymbol{s},(n,q)} = \boldsymbol{0}_{L_{\mathrm{w}} \times L_{\mathrm{a}}}$ if $q \neq \chi_n(m, \boldsymbol{s})$

B. Conditions of Achieving the PIR Capacity

This subsection discuss the capacity-achieving conditions using its extended parity check matrix. Specifically, we discuss the rest $\tilde{\boldsymbol{H}}'$ in Condition 3.1. We first prepare lemmas and notations.

Lemma 3.3: A LPIR achieves the PIR capacity iff the LPIR satisfies Condition 3.3.

Condition 3.3 (Capacity-Achieving Conditions Using its Generator Matrix):

(1) For any $m \in [M], \boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}, n, n' \in [N]$,

$$\mathrm{rank}\left(\boldsymbol{L}^{(m)}, \boldsymbol{G}_{(n,q_n)}, \boldsymbol{G}_{(n',q_{n'})}\right) - \mathrm{rank}\left(\boldsymbol{L}^{(m)}, \boldsymbol{G}_{(n,q_n)}\right) = 0 \tag{15}$$

(2) For any $m \in [M], \boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}, I \subset [M] \setminus \{m\}$ with $I \neq \emptyset$,

$$\mathrm{rank}\left(\boldsymbol{L}^I, \boldsymbol{G}_{R(\boldsymbol{q})}\right) - |I|L_{\mathrm{w}} = \sum_{n \in [N]} \left(\mathrm{rank}\left(\boldsymbol{L}^I, \boldsymbol{G}_{(n,q_n)}\right) - |I|L_{\mathrm{w}}\right). \tag{16}$$

(3) $\sum_{\boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(1)}} \mathrm{rank}\boldsymbol{G}_{R(\boldsymbol{q})} = \cdots = \sum_{\boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(M)}} \mathrm{rank}\boldsymbol{G}_{R(\boldsymbol{q})}$,

where $\boldsymbol{L}^{(m)} := \begin{pmatrix} \boldsymbol{0}_{(m-1)L_{\mathrm{w}} \times L_{\mathrm{w}}} \\ \boldsymbol{I}_{L_{\mathrm{w}}} \\ \boldsymbol{0}_{(M-m)L_{\mathrm{w}} \times L_{\mathrm{w}}} \end{pmatrix}$ and $\boldsymbol{L}^I := \left(\boldsymbol{L}^{(m)} | m \in I\right) \in \mathbb{F}_p^{ML_{\mathrm{w}} \times |I| L_{\mathrm{w}}}$.

The proof is in Appendix B in [1].

Condition 3.3 is equivalent to Condition 10 in [8]. On the other hand, Condition 3.3 is slightly different from Conditions P1,P2,P3 in [4] because the definitions of the rate are different.

Hereafter we rewrite (1), (2) and (3) of Condition 3.3 by the EPCM $\tilde{\boldsymbol{H}}$ instead of $\boldsymbol{G}$, respectively.

Lemma 3.4: (1) of Condition 3.3 is a necessary and sufficient condition for the following condition.

Condition 3.4: $\tilde{\boldsymbol{H}}'$ in Condition 3.1 has the form of $\tilde{\boldsymbol{H}}' = \begin{pmatrix} \boldsymbol{K} \\ \vdots \end{pmatrix}$, where $\boldsymbol{K}$ has the form in Fig. 1. We abbriviate $A := \sum_{n \in [N]} |\mathcal{Q}_n|$, $C := (M-1)A(A-1)$ in Fig. 1. Moreover, the following (i) and (ii) hold in Fig. 1.

(i) Each block matrix in the first M block columns is a $L_{\mathrm{a}} \times L_{\mathrm{w}}$ submatrix. Each $\tilde{\boldsymbol{H}}_{i,j}$ is any matrix.

(ii) Each block matrix in the last A block columns is a $L_{\mathrm{a}} \times L_{\mathrm{a}}$ submatrix. Each $\tilde{\boldsymbol{H}}_{i,j}$ is any matrix.

Proof : (1) of Condition 3.3 is equivalent to that for any $m, n, n, \boldsymbol{q}$, each column of $\boldsymbol{G}_{(n',q_{n'})}$ is a linear combination

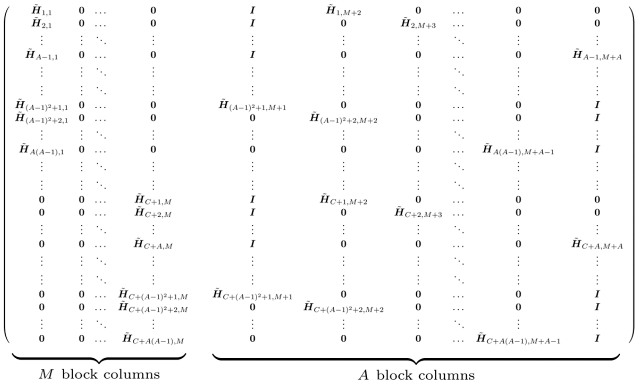

Fig. 1. The figure of $\boldsymbol{K}$

of all columns of $\boldsymbol{L}^{(m)}, \boldsymbol{G}_{(n,q_n)}, \boldsymbol{G}_{(n',q_{n'})}$. This is equivalent to $\tilde{\boldsymbol{G}}\boldsymbol{K}^\top = \boldsymbol{0}$. □

Definition 3.3: Let $\boldsymbol{G}_{(n,q)} = (\boldsymbol{g}_{(n,q),1}, \dots, \boldsymbol{g}_{(n,q),L_\mathrm{a}})$ be the column vector representation of a submatrix $\boldsymbol{G}_{(n,q)}$ of the generator matrix $\boldsymbol{G}$ of an LPIR. Let $\boldsymbol{l}_i$ be the i th column of $\boldsymbol{L}^{[M]}$ for any $i \in [M] \times [L_\mathrm{a}]$, i.e. $\boldsymbol{L}^{[M]} = \{\, \boldsymbol{l}_i \mid i \in [M] \times [L_\mathrm{a}] \,\}$. We abbriviate $\mathcal{NQ} := \{\, (n,q) \mid n \in [N], q \in \mathcal{Q}_n \,\}$. We fix a function $B(\cdot,\cdot)\colon 2^{[M]} \times 2^{\mathcal{NQ}} \to 2^{\mathcal{NQ}\times[L_\mathrm{a}]}$; $(I,R) \mapsto B(I,R)$ satisfying the following conditions.

(1) The union set of a set $\{\, \boldsymbol{l}_i \mid i \in I \times [L_\mathrm{a}] \,\}$ of the columns of $\boldsymbol{G}_R$ and a set $\left\{\, \boldsymbol{g}_{(n,q),j} \,\middle|\, ((n,q),j) \in B(I,R) \,\right\}$ of the columns of $\boldsymbol{G}_R$ is a basis of a linear code generated by all columns of a matrix $\left(\boldsymbol{L}^I, \boldsymbol{G}_R\right)$ for any $I \subset [M]$ and $R \subset \mathcal{NQ}$.

(2) $B(I,\emptyset) := \emptyset$ for any $I \subset [M]$.

(3) For any $I \subset [M]$ and $R \subset \mathcal{NQ}$, it holds that $B(I,R) \subset R \times [L_a]$.

(4) We exhaustively partition a set $R \subset \mathcal{NQ}$ into $R = R_1 \sqcup R_2$ satisfying $R_1, R_2 \neq \emptyset$. Then for any $I \subset [M]$, it holds that $B(I,R) \subset B(I,R_1) \cup B(I,R_2)$.

(5) For any $I_2 \subset I_1 \subset [M]$ and $R \subset \mathcal{NQ}$, it holds that $B(I_2,R) \subset B(I_1,R)$.

For an LPIR, the function $B(\cdot,\cdot)$ satisfying the above is not necessarily uniquely determined. We fix any one of them.

Lemma 3.5: All column vectors in $\{\, \boldsymbol{l}_i \mid i \in I \times [L_\mathrm{a}] \,\} \cup \left\{\, \boldsymbol{g}_{(n,q),j} \,\middle|\, (n,q,i) \in B(I,R(\boldsymbol{q})) \,\right\}$ generate each column vector in $\left\{\, \boldsymbol{g}_{(n,q),j} \,\middle|\, (n,q,i) \notin B(I,R(\boldsymbol{q})) \,\right\}$ iff Condition 3.5 holds.

Condition 3.5: $\tilde{\boldsymbol{H}}'$ in Condition 3.1 has the form of $\tilde{\boldsymbol{H}}' = \begin{pmatrix} \boldsymbol{K} \\ \boldsymbol{B} \\ \vdots \end{pmatrix}$, where $\boldsymbol{K}$ is in Condition 3.4. $\boldsymbol{B}$ is a submatrix that collects matrices of the form shown in Fig. 2 and 3 for any $m \in [M], I := [M] \setminus \{m\}, \boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}, n \in [N]$.

The proof of Lemma 3.5 is in Appendix C in [1].

Lemma 3.6: All vectors in $\{\, \boldsymbol{l}_i \mid i \in I \times [L_\mathrm{a}] \,\} \cup \left\{\, \boldsymbol{g}_{(n,q),j} \,\middle|\, (n,q,i) \in B(I,R(\boldsymbol{q})) \,\right\}$ are linearly independent iff Condition 3.6 holds.

Condition 3.6: For any $m \in [M]$,$\boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}$, $I \subset [M]\setminus\{m\}$, the following conditions hold.

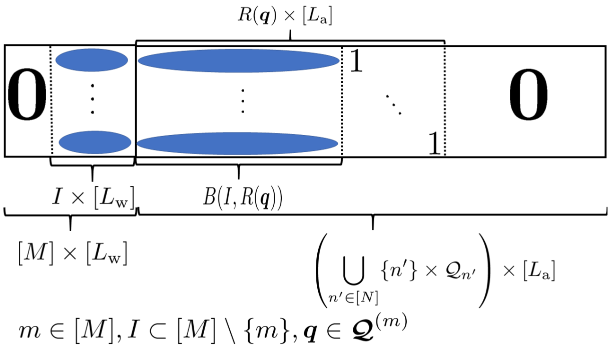
$R(\boldsymbol{q}) \times [L_\mathrm{a}]$
$\mathbf{0}$
1
1
$\mathbf{0}$
$I \times [L_\mathrm{w}]$
$B(I,R(\boldsymbol{q}))$
$[M] \times [L_\mathrm{w}]$
$\left(\bigcup_{n'\in[N]} \{n'\} \times \mathcal{Q}_{n'}\right) \times [L_\mathrm{a}]$
$m \in [M], I \subset [M] \setminus \{m\}, \boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}$

Fig. 2. A part of block rows of the EPCM $\tilde{\boldsymbol{H}}$

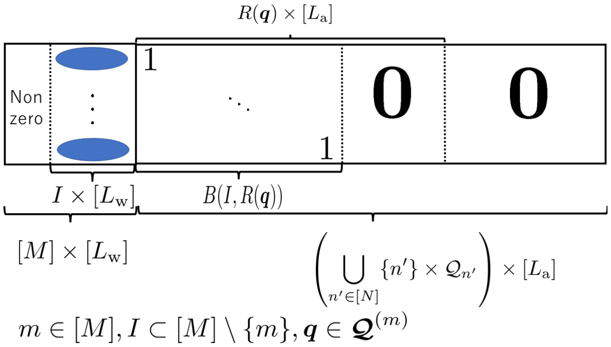
$R(\boldsymbol{q}) \times [L_\mathrm{a}]$
Non zero
1
1
$\mathbf{0}$
$\mathbf{0}$
$I \times [L_\mathrm{w}]$
$B(I,R(\boldsymbol{q}))$
$[M] \times [L_\mathrm{w}]$
$\left(\bigcup_{n'\in[N]} \{n'\} \times \mathcal{Q}_{n'}\right) \times [L_\mathrm{a}]$
$m \in [M], I \subset [M] \setminus \{m\}, \boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}$

Fig. 3. A part of block rows of the EPCM $\tilde{\boldsymbol{H}}$

(1) no vector is generated whose support is a subset of $(I \times [L_\mathrm{w}]) \sqcup B(I, R(\boldsymbol{q}))$ (the blue set in each row in Fig.4, where $R = R(\boldsymbol{q})$) by all transpose of rows of the EPCM $\tilde{\boldsymbol{H}}$.

(2) Let $I \neq \emptyset$. For any $n \in [N]$, no vector is generated whose support is a subset of $(I \times [L_\mathrm{w}]) \sqcup B(I, \{(n,q_n)\})$ by all transpose of rows of the EPCM $\tilde{\boldsymbol{H}}$.

(3) No vector is generated whose support is a subset of $[M] \times [L_\mathrm{w}]$ by all transpose of rows of the EPCM $\tilde{\boldsymbol{H}}$.

Proof : (1) holds because for any $m, I, \boldsymbol{q}$, all columns of the union of the sets $\{\, \boldsymbol{l}_i \mid i \in I \times [L_\mathrm{w}] \,\}$ and $\left\{\, \boldsymbol{g}_{(n,q),j} \,\middle|\, ((n,q),j) \in B(I,R(\boldsymbol{q})) \,\right\}$ are linearly independent. (2) is similar. (3) holds because the matrix corresponding to the index set in the EGM $\boldsymbol{G}$ is $\boldsymbol{I}_{ML_\mathrm{w}}$. □

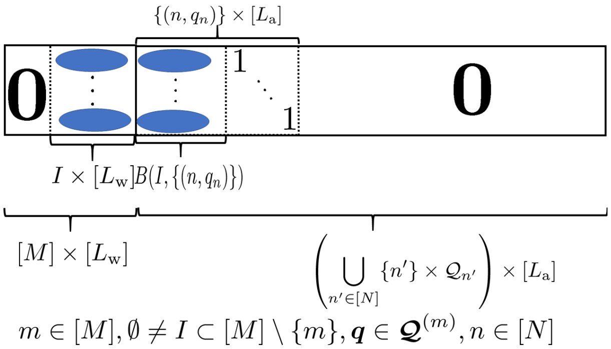
$\{(n,q_n)\} \times [L_\mathrm{a}]$
$\mathbf{0}$
1
1
$\mathbf{0}$
$I \times [L_\mathrm{w}]$
$B(I,\{(n,q_n)\})$
$[M] \times [L_\mathrm{w}]$
$\left(\bigcup_{n'\in[N]} \{n'\} \times \mathcal{Q}_{n'}\right) \times [L_\mathrm{a}]$
$m \in [M], \emptyset \neq I \subset [M] \setminus \{m\}, \boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}, n \in [N]$

Fig. 4. A part of block rows of the EPCM $\tilde{\boldsymbol{H}}$

Lemma 3.7: (2) of Condition 3.3 is a necessary and sufficient condition for the following condition.

Condition 3.7: For any $m \in [M], \boldsymbol{q} = (q_1, \ldots, q_N) \in \boldsymbol{\mathcal{Q}}^{(m)}, I \subset [M] \setminus \{m\}$ with $I \neq \emptyset$,

$$B\left(I, R\left(\boldsymbol{q}\right)\right) = \bigcup_{n \in [N]} B\left(I, \{\ (n, q_n)\ \}\right). \tag{17}$$

We choose to use a direct sum in the R.H.S. since they are mutually exclusive althogh the R.H.S could be a union set.

Proof : Assume that (2) of Condition 3.3 holds. We show that Eq.(17) holds. $B\left(I, R\left(\boldsymbol{q}\right)\right) \subset \bigcup_{n \in [N]} B\left(I, \{\ (n, q_n)\ \}\right)$. For any $m \in [M], \boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}, I \subset [M] \setminus \{m\}$, $I \neq \emptyset$,

$$\begin{aligned} 0 &= \operatorname{rank}\left(\boldsymbol{L}^I, \boldsymbol{G}_{(1,q_1),\ldots,(N,q_N)}\right) - \sum_{n \in [N]} \operatorname{rank}\left(\boldsymbol{L}^I, \boldsymbol{G}_{(n,q_n)}\right) \\ &\quad + (N-1)|I|L_{\mathrm{w}} \qquad (18) \\ &= |I|L_{\mathrm{w}} + \left|B\left(I, R\left(\boldsymbol{q}\right)\right)\right| - \sum_{n \in [N]} \left(|I|L_{\mathrm{w}} + \left|B\left(I, \{\ (n, q_n)\ \}\right)\right|\right) \\ &\quad + (N-1)|I|L_{\mathrm{w}} \qquad (19) \\ &= \left|B\left(I, R\left(\boldsymbol{q}\right)\right)\right| - \sum_{n \in [N]} \left|B\left(I, \{\ (n, q_n)\ \}\right)\right|. \qquad (20) \end{aligned}$$

The converse is clear from the above equations. □

Lemma 3.8: (3) of Condition 3.3 is a necessary and sufficient condition for the following condition.

Condition 3.8:

$$\sum_{\boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(1)}} \left|B\left(\emptyset, R\left(\boldsymbol{q}\right)\right)\right| = \cdots = \sum_{\boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(M)}} \left|B\left(\emptyset, R\left(\boldsymbol{q}\right)\right)\right| \tag{21}$$

Proof : $B\left(\emptyset, R\left(\boldsymbol{q}\right)\right)$ is the basis of the linear code generated by all columns of $\boldsymbol{G}_{R(\boldsymbol{q})}$.□

Theorem 3.2: If there exists a function χ, a function B and a matrix $\tilde{\boldsymbol{H}} = \begin{pmatrix} -\boldsymbol{J} & \tilde{\boldsymbol{D}} \\ & \tilde{\boldsymbol{H}}' \end{pmatrix}$, where $\tilde{\boldsymbol{H}}' = \begin{pmatrix} \boldsymbol{K} \\ \boldsymbol{B} \\ \boldsymbol{C} \end{pmatrix}$ and $\boldsymbol{C} \in \mathbb{F}_p^{h'' \times (L_{\mathrm{a}} \sum_{n \in [N]} |\mathcal{Q}_n| + M L_{\mathrm{w}})}$ for some $h'' \geq 0$, satisfying from Conditions 3.2, 3.4, 3.5, 3.6, 3.7 and 3.8, we can construct a CALPIR whose generator matrix is $\boldsymbol{G}$ in Condition 3.1.

This theorem makes the condition for the generating matrix $\boldsymbol{G}$ to constitute a large subclass of CALPIRs more explicit than Condition 3.3.

IV. Conclusion

This study shows several conditions of a CALPIR using its EPCM. The future work is to propose a systematic construction method for the EPCM $\tilde{\boldsymbol{H}}$, especially $\boldsymbol{C}$.

REFERENCES

[1] K. Kazama and T Yoshida. A note on parity check matrix of private information retrieval code. https://x.gd/kEJ58, 2024.

[2] Benny Chor, Eyal Kushilevitz, Oded Goldreich, and Madhu Sudan. Private information retrieval. *J. ACM*, Vol. 45, No. 6, pp. 965–981, November 1998.

[3] H. Sun and S. A. Jafar. The capacity of private information retrieval. *IEEE Transactions on Information Theory*, Vol. 63, No. 7, pp. 4075–4088, 2017.

[4] C. Tian, H. Sun, and J. Chen. Capacity-achieving private information retrieval codes with optimal message size and upload cost. In *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, pp. 1–6, May 2019.

[5] Siddhartha Kumar, Hsuan-Yin Lin, Eirik Rosnes, and Alexandre Graell i Amat. Achieving private information retrieval capacity in distributed storage using an arbitrary linear code. *CoRR*, Vol. abs/1712.03898, , 2017.

[6] Ragnar Freij-Hollanti, Oliver W. Gnilke, Camilla Hollanti, Anna-Lena Horlemann-Trautmann, David A. Karpuk, and Ivo Kubjas. t-private information retrieval schemes using transitive codes. *CoRR*, Vol. abs/1712.02850, , 2017.

[7] Hua Sun and Syed Ali Jafar. On the capacity of locally decodable codes. *IEEE Transactions on Information Theory*, Vol. 66, No. 10, pp. 6566–6579, 2020.

[8] U. Imazu, K. Kazama, and T. Matsushima. A proposal of the construction algorithm of private information retrieval systems with high efficiency (in japanese). *Proceedings of Japan Industrial Management Association Annual Conference 2021 Spring*, pp. 367–370, May 2021.

[9] K. Kazama, A. Kamatsuka, T. Yoshida, and T. Matsushima. A note on a relationship between smooth locally decodable codes and private information retrieval. In *2020 International Symposium on Information Theory and Its Applications (ISITA)*, pp. 259–263, 2020.

APPENDIX

A. *The Proof of the Iniquality* (13)

Proof :

$$\frac{L_{\mathrm{w}}}{\mathrm{r_I}} = \max_{m\in[M]} \mathrm{H}\left(\boldsymbol{A}^{(m)}|\boldsymbol{Q}^{(m)}\right) \tag{22}$$

$$\overset{\text{(d)}}{\geq} \mathrm{H}\left(\boldsymbol{A}^{(1)}|\boldsymbol{Q}^{(1)}\right) \tag{23}$$

$$= \sum_{\boldsymbol{q}\in\boldsymbol{\mathcal{Q}}^{(1)}} \Pr(\boldsymbol{Q}^{(1)}=\boldsymbol{q})\mathrm{H}(\Phi(\boldsymbol{W})_{R(\boldsymbol{Q}^{(1)})}|\boldsymbol{Q}^{(1)}=\boldsymbol{q}) \tag{24}$$

$$= \sum_{\boldsymbol{q}\in\boldsymbol{\mathcal{Q}}^{(1)}} \Pr(\boldsymbol{Q}^{(1)}=\boldsymbol{q})\mathrm{H}(\Phi(\boldsymbol{W})_{R(\boldsymbol{q})}) \tag{25}$$

$$= \sum_{\boldsymbol{q}\in\boldsymbol{\mathcal{Q}}^{(1)}} \Pr(\boldsymbol{Q}^{(1)}=\boldsymbol{q})\left(\mathrm{H}(\Phi(\boldsymbol{W})_{R(\boldsymbol{q})}|W_1)+\mathrm{H}(W_1)\right) \tag{26}$$

$$= L_{\mathrm{w}} + \sum_{\boldsymbol{q}\in\boldsymbol{\mathcal{Q}}^{(1)}} \Pr(\boldsymbol{Q}^{(1)}=\boldsymbol{q})\mathrm{H}(\Phi(\boldsymbol{W})_{R(\boldsymbol{q})}|W_1). \tag{27}$$

For any $m \in [2, M]$,

$$\sum_{\boldsymbol{q}\in\boldsymbol{\mathcal{Q}}^{(m-1)}} \Pr(\boldsymbol{Q}^{(m-1)}=\boldsymbol{q})\mathrm{H}(\Phi(\boldsymbol{W})_{R(\boldsymbol{q})}|\boldsymbol{W}_{[m-1]}) \tag{28}$$

$$\overset{\text{(a)}}{\geq} \sum_{\boldsymbol{q}\in\boldsymbol{\mathcal{Q}}^{(m-1)}} \Pr(\boldsymbol{Q}^{(m-1)}=\boldsymbol{q})\frac{1}{N} \times \sum_{n\in[N]} \mathrm{H}(\Phi(\boldsymbol{W})_{(n,q_n)}|\boldsymbol{W}_{[m-1]}) \tag{29}$$

$$= \frac{1}{N}\sum_{n\in[N]}\sum_{q\in\mathcal{Q}_n}\left(\sum_{\substack{\boldsymbol{q}\in\boldsymbol{\mathcal{Q}}^{(m-1)}\\ q_n=q}} \Pr(\boldsymbol{Q}^{(m-1)}=\boldsymbol{q})\right) \times \mathrm{H}(\Phi(\boldsymbol{W})_{(n,q)}|\boldsymbol{W}_{[m-1]}) \tag{30}$$

$$= \frac{1}{N}\sum_{n\in[N]}\sum_{q\in\mathcal{Q}_n}\left(\sum_{\substack{\boldsymbol{q}\in\boldsymbol{\mathcal{Q}}^{(m)}\\ q_n=q}} \Pr(\boldsymbol{Q}^{(m)}=\boldsymbol{q})\right) \times \mathrm{H}(\Phi(\boldsymbol{W})_{(n,q)}|\boldsymbol{W}_{[m-1]}) \quad \because \text{Privacy constraint} \tag{31}$$

$$= \sum_{\boldsymbol{q}\in\boldsymbol{\mathcal{Q}}^{(m)}} \Pr(\boldsymbol{Q}^{(m)}=\boldsymbol{q})\frac{1}{N}\sum_{n\in[N]}\mathrm{H}(\Phi(\boldsymbol{W})_{(n,q_n)}|\boldsymbol{W}_{[m-1]}) \tag{32}$$

$$\overset{\text{(b)}}{\geq} \sum_{\boldsymbol{q}\in\boldsymbol{\mathcal{Q}}^{(m)}} \Pr(\boldsymbol{Q}^{(m)}=\boldsymbol{q})\frac{1}{N}\mathrm{H}(\Phi(\boldsymbol{W})_{R(\boldsymbol{q})}|\boldsymbol{W}_{[m-1]}) \tag{33}$$

$$= \sum_{\boldsymbol{q}\in\boldsymbol{\mathcal{Q}}^{(m)}} \frac{1}{N}\Pr(\boldsymbol{Q}^{(m)}=\boldsymbol{q})\left(\mathrm{H}(\Phi(\boldsymbol{W})_{R(\boldsymbol{q})}|\boldsymbol{W}_{[m]})+\mathrm{H}(W_m)\right) \tag{34}$$

$$= \frac{L_{\mathrm{w}}}{N}+\frac{1}{N}\left(\sum_{\boldsymbol{q}\in\boldsymbol{\mathcal{Q}}^{(m)}}\Pr\left(\boldsymbol{Q}^{(m)}=\boldsymbol{q}\right)\mathrm{H}(\Phi(\boldsymbol{W})_{R(\boldsymbol{q})}|\boldsymbol{W}_{[m]})\right). \tag{35}$$

Thus

$$\begin{aligned}
\frac{L_{\mathrm{w}}}{\mathrm{r_I}} &\geq L_{\mathrm{w}} + \sum_{\boldsymbol{q}\in\boldsymbol{\mathcal{Q}}^{(1)}} \Pr(\boldsymbol{Q}^{(1)}=\boldsymbol{q})\mathrm{H}(\Phi(\boldsymbol{W})_{R(\boldsymbol{q})}|W_1)\\
&\geq L_{\mathrm{w}}\left(1+\frac{1}{N}\right)\\
&\quad+\frac{1}{N}\sum_{\boldsymbol{q}\in\boldsymbol{\mathcal{Q}}^{(2)}}\Pr(\boldsymbol{Q}^{(2)}=\boldsymbol{q})\mathrm{H}(\Phi(\boldsymbol{W})_{R(\boldsymbol{q})}|\boldsymbol{W}_{[2]})\\
&\cdots\\
&\geq L_{\mathrm{w}}\left(1+\frac{1}{N}+\cdots+\frac{1}{N^{M-2}}\right)+\frac{1}{N^{M-2}}\sum_{\boldsymbol{q}\in\boldsymbol{\mathcal{Q}}^{(M-1)}}\\
&\quad\times\Pr(\boldsymbol{Q}^{(M-1)}=\boldsymbol{q})\mathrm{H}(\Phi(\boldsymbol{W})_{R(\boldsymbol{q})}|\boldsymbol{W}_{[M-1]})\\
&\geq L_{\mathrm{w}}\left(1+\frac{1}{N}+\cdots+\frac{1}{N^{M-2}}+\frac{1}{N^{M-1}}\right)+\frac{1}{N^{M-1}}\\
&\quad\times\sum_{\boldsymbol{q}\in\boldsymbol{\mathcal{Q}}^{(M)}}\Pr\left(\boldsymbol{Q}^{(M)}=\boldsymbol{q}\right)\mathrm{H}(\Phi(\boldsymbol{W})_{R(\boldsymbol{q})}|\boldsymbol{W}_{[M]})\\
&= L_{\mathrm{w}}\left(1+\frac{1}{N}+\cdots+\frac{1}{N^{M-1}}\right).
\end{aligned}$$

□

B. *The Proof of Lemma 3.3*

Proof : $\overset{\text{(a)}}{\leq}$, $\overset{\text{(b)}}{\leq}$, $\overset{\text{(c)}}{\leq}$ and $\overset{\text{(d)}}{\leq}$ in Appendix A shows that a PIR achieves the PIR capacity iff[1]

[1] We can replace Eq.(37) with $\mathrm{H}\left(\Phi(\boldsymbol{W})_{(1,q_1)},\ldots,\Phi(\boldsymbol{W})_{(N,q_N)},\boldsymbol{W}_I\right) - \mathrm{H}\left(\boldsymbol{W}_I\right) = \sum_{n\in[N]}\mathrm{H}\left(\left(\Phi(\boldsymbol{W})_{(n,q_n)},\boldsymbol{W}_I\right)-\mathrm{H}\left(\boldsymbol{W}_I\right)\right)$ and Eq.(36) with $\mathrm{H}\left(\Phi(\boldsymbol{W})_{(n,q_n)},\Phi(\boldsymbol{W})_{(n',q_{n'})},W_m\right)-\mathrm{H}\left(\Phi(\boldsymbol{W})_{(n',q_{n'})},W_m\right)=0.$

(1) For any $m \in [M], \boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}, n, n' \in [N]$,

$$\mathrm{H}\left(\Phi(\boldsymbol{W})_{(n,q_n)} | \Phi(\boldsymbol{W})_{(n',q_{n'})}, W_m\right) = 0. \tag{36}$$

(2) For any $m \in [M], \boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}, I \subset [M] \setminus \{m\}$ with $I \neq \emptyset$,

$$\begin{aligned} &\mathrm{H}\left(\Phi(\boldsymbol{W})_{(1,q_1)}, \ldots, \Phi(\boldsymbol{W})_{(N,q_N)} | \boldsymbol{W}_I\right) \\ &= \sum_{n \in [N]} \mathrm{H}\left(\Phi(\boldsymbol{W})_{(n,q_n)} | \boldsymbol{W}_I\right). \end{aligned} \tag{37}$$

(3)

$$\sum_{\boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(1)}} \mathrm{H}\left(\Phi(\boldsymbol{W})_{R(\boldsymbol{q})}\right) = \cdots = \sum_{\boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(M)}} \mathrm{H}\left(\Phi(\boldsymbol{W})_{R(\boldsymbol{q})}\right). \tag{38}$$

These are equivalent to Condition 3.3.□

Remark A.1: If an IR is an LIR, the rate is

$$\mathrm{rate}_{\mathrm{I}} = \min_{m \in [M]} \frac{L_{\mathrm{w}} S}{\sum_{\boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}} \operatorname{rank}\left(\boldsymbol{G}_{R(\boldsymbol{q})}\right)}. \tag{39}$$

C. The Proof of Lemma 3.5

Proof : We focus on Fig.4,5,2 and 3.

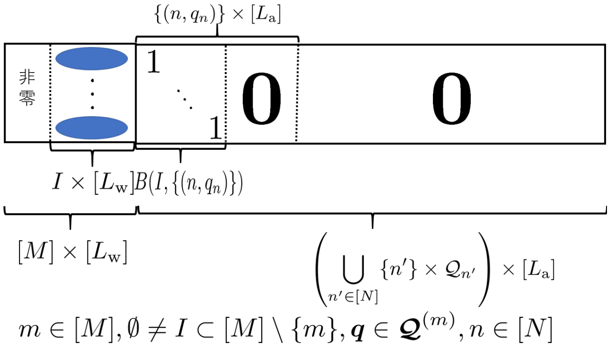
$\{(n, q_n)\} \times [L_{\mathrm{a}}]$
非零
1
1
0
0
$I \times [L_{\mathrm{w}}]$
$B(I, \{(n, q_n)\})$
$[M] \times [L_{\mathrm{w}}]$
$\left(\bigcup_{n' \in [N]} \{n'\} \times \mathcal{Q}_{n'}\right) \times [L_{\mathrm{a}}]$
$m \in [M], \emptyset \neq I \subset [M] \setminus \{m\}, \boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}, n \in [N]$

Fig. 5. A part of block rows of the EPCM $\tilde{\boldsymbol{H}}$

- $m \in [M]$, $I \subset [M] \setminus \{m\}$,$\boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}$ を任意にとる．行列 $\boldsymbol{L}^I$ の列ベクトル全体の集合 $\{\, \boldsymbol{l}_i \mid i \in I \times [L_{\mathrm{a}}] \,\}$ と集合 $\left\{\, \boldsymbol{g}_{(n'',q''),i} \,\middle|\, ((n'', q''), i) \in B\left(I, R(\boldsymbol{q})\right) \,\right\}$ の和集合は行列 $\left(\boldsymbol{L}^I, \boldsymbol{G}_{R(\boldsymbol{q})}\right)$ における他の列ベクトル集合を生成するので，拡大検査行列 $\tilde{\boldsymbol{H}}$ の一部の行が，図 2 の行列 $\in \mathbb{F}_p^{(L_{\mathrm{a}}|R(\boldsymbol{q})| - B(I,R(\boldsymbol{q}))) \times (ML_{\mathrm{w}} + \sum_{n \in [N]} |\mathcal{Q}_n| L_{\mathrm{a}})}$ である．ただし，左の $\boldsymbol{0}$ は $(L_{\mathrm{a}}|R(\boldsymbol{q})| - B(I, R(\boldsymbol{q}))) \times (ML_{\mathrm{w}} - |I|\, L_{\mathrm{w}})$ 零行列であり，右の $\boldsymbol{0}$ は $(L_{\mathrm{a}}|R(\boldsymbol{q})| - B(I, R(\boldsymbol{q}))) \times (-1 + \sum_{n \in [N]} |\mathcal{Q}_n|) L_{\mathrm{a}}$ 零行列である．(図の青の部分の成分はすべて 0 でも構わないとする)
  また，$\boldsymbol{G}_{R(\boldsymbol{q})}$ の列ベクトル集合 $\left\{\, \boldsymbol{g}_{(n'',q''),i} \,\middle|\, ((n'', q''), i) \in B\left(I, R(\boldsymbol{q})\right) \,\right\}$ と行列 $\boldsymbol{L}^I$ の全列ベクトルは一次独立であるが，列ベクトル集合 $\left\{\, \boldsymbol{g}_{(n'',q''),i} \,\middle|\, ((n'', q''), i) \in B\left(I, R(\boldsymbol{q})\right) \,\right\}$ は行列 $\boldsymbol{L}$ から生成されるので，拡大検査行列 $\tilde{\boldsymbol{H}}$ の一部の行が図 3 の形の行列 $\in \mathbb{F}_p^{|B(I,R(\boldsymbol{q}))| \times (ML_{\mathrm{w}} + \sum_{n \in [N]} |\mathcal{Q}_n| L_{\mathrm{a}})}$ である．
- $m \in [M]$, $(\emptyset \neq) I \subset [M] \setminus \{m\}$,$\boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}, n \in [N]$ を任意にとる．
  先と同様に，拡大検査行列 $\tilde{\boldsymbol{H}}$ の一部の行が図 4,5 の形の行列になる．

ここで，行列 $\boldsymbol{B}'$ を，各 $m \in [M], I \subset [M] \setminus \{m\}, \boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}$ に対する図 4,5 の形の行列および各 $m \in [M], \emptyset \neq I \subset [M] \setminus \{m\}, \boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}, n \in [N]$ に対する図 2,3 の形の行列を集めた部分行列とおく．簡約化して零行列を取り除くと行列 $\boldsymbol{B}$ の形になることを示す．

今，$R \subset \left(\bigcup_{n \in [N]} \{n\} \times \mathcal{Q}_n\right) \times [L_{\mathrm{a}}]$ を任意に固定して，$I \subset [M]$ を任意に動かした時，I の包含順に $B(I, R)$ から作られる図 6 の形の行列を並べると図 7 のようになる．

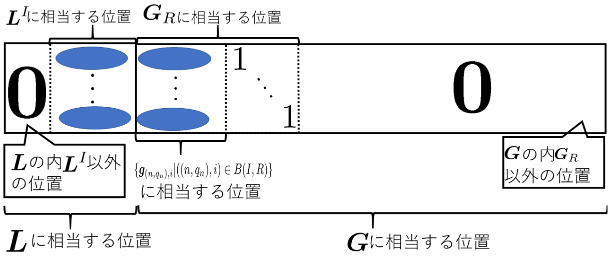
$\boldsymbol{L}^I$に相当する位置
$\boldsymbol{G}_R$に相当する位置
0
1
1
0
$\boldsymbol{L}$の内$\boldsymbol{L}^I$以外の位置
$\{\boldsymbol{g}_{(n,q_n),i} | ((n, q_n), i) \in B(I, R)\}$ に相当する位置
$\boldsymbol{G}$の内$\boldsymbol{G}_R$以外の位置
$\boldsymbol{L}$に相当する位置
$\boldsymbol{G}$に相当する位置

Fig. 6. A part of block rows of the EPCM $\tilde{\boldsymbol{H}}$

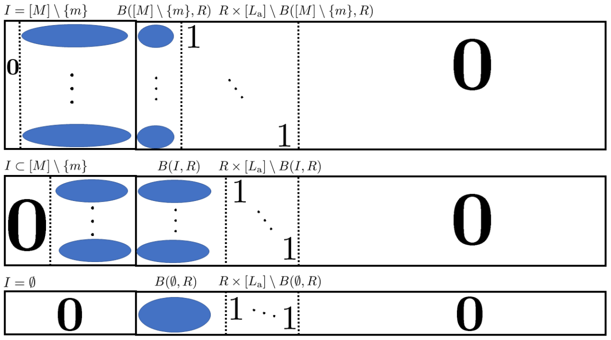
$I = [M] \setminus \{m\}$
$B([M] \setminus \{m\}, R)$
$R \times [L_{\mathrm{a}}] \setminus B([M] \setminus \{m\}, R)$
0
1
1
0
$I \subset [M] \setminus \{m\}$
$B(I, R)$
$R \times [L_{\mathrm{a}}] \setminus B(I, R)$
0
1
1
0
$I = \emptyset$
$B(\emptyset, R)$
$R \times [L_{\mathrm{a}}] \setminus B(\emptyset, R)$
0
1 ⋯ 1
0

Fig. 7. A part of block rows of the EPCM $\tilde{\boldsymbol{H}}$

この関係と後述の箇条書き**??**番目から，各 $m \in [M], I \subset [M] \setminus \{m\}, \boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}$ に対し図 4 の形の行列を集めた部分行列は，簡約化によって，各 $m \in [M], I := [M] \setminus \{m\}, \boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}$ に対し図 4 の形の行列を集めた部分行列によって消滅する．図 2 についても同様のことが成立する．さらに，詳細な説明は省くが，図 5,3 についても同様のことが成立する．

以上から，行列 $\boldsymbol{B}'$ のうち，各 $m \in [M], I := [M] \setminus \{m\}, \boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}$ に対する図 4,5 の形の行列，および各 $m \in [M], I := [M] \setminus \{m\}, \boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}, n \in [N]$ に対する図 2,3 の形の行列を集めた部分行列 $\boldsymbol{B}''$ のみを考えればよいということになる．

ここで，各 $m \in [M], I := [M] \setminus \{m\}, \boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}$ に対し，図 4,5 の形の行列は，各 $n \in [N]$ に対し図 2,3 の形の行列を集めてできた部分行列によって簡約化されることにより，消滅する．

図 4 の形の行列は，各 $n \in [N]$ に対する図 2 の形の行列を集めた部分行列によって簡約化されることにより，消滅する

ことについて説明する．後者を全射によって簡約化すると，図 8 のような行列に変形される．しかし，これは後述の**??**,**??**

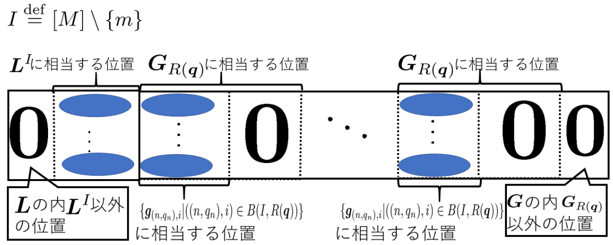
$I \overset{\text{def}}{=} [M] \setminus \{m\}$
$\boldsymbol{L}^I$に相当する位置
$\boldsymbol{G}_{R(\boldsymbol{q})}$に相当する位置
$\boldsymbol{G}_{R(\boldsymbol{q})}$に相当する位置
$\mathbf{0}$
$\mathbf{0}$
$\mathbf{0}$
$\mathbf{0}$
$\boldsymbol{L}$の内$\boldsymbol{L}^I$以外の位置
$\{\boldsymbol{g}_{(n,q_n),i} | ((n,q_n),i) \in B(I,R(\boldsymbol{q}))\}$に相当する位置
$\{\boldsymbol{g}_{(n,q_n),i} | ((n,q_n),i) \in B(I,R(\boldsymbol{q}))\}$に相当する位置
$\boldsymbol{G}$の内$\boldsymbol{G}_{R(\boldsymbol{q})}$以外の位置

Fig. 8. A part of block rows of the EPCM $\tilde{\boldsymbol{H}}$

番目の条件により，すべての行は零になる．よって，行列 $\boldsymbol{B}''$ は，各 $m \in [M], I \coloneqq [M] \setminus \{m\}, \boldsymbol{q} \in \boldsymbol{\mathcal{Q}}^{(m)}, n \in [N]$ に対し図 2,3 の形の行列を集めた部分行列 $\boldsymbol{B}$ へと変形される．

図 5 の行列が，各 $n \in [N]$ に対し図 3 の形の行列を集めてできた部分行列によって消滅することも同様に示される．

以上を合わせると，拡大検査行列の一部は部分行列 $\boldsymbol{B}$ の形になる．□