# On the Linear Capacity of Conditional Disclosure of Secrets

Zhou Li and Hua Sun
Department of Electrical Engineering
University of North Texas, Denton, TX 76203
*Email: zhouli@my.unt.edu, hua.sun@unt.edu*

***Abstract*—Conditional disclosure of secrets (C**
**lem of disclosing as efficiently as possible, one**
**and Bob to Carol if and only if the inputs a**
**satisfy some function $f$. The information theo**
**CDS is the maximum number of bits of the**
**be securely disclosed per bit of total commun**
**instances, where the capacity is the highest and**
**are recently characterized through a noise and**
**approach and are described using a graph repr**
**function $f$, $G_f$. In this work, we go beyond the b**
**and further develop the alignment approach to**
**linear capacity of a class of CDS instances to**
**where $\rho$ is a covering parameter of $G_f$.**

## I. Introduction

The conditional disclosure of secrets (CD
classical cryptographic primitive[1] with rich
many other primitives such as symmetric pri
retrieval [2] and secret sharing [3], [4]. The
problem is to find the most efficient way for
disclose a common secret to Carol if and only if the inputs at Alice and Bob satisfy some function $f$ (see Fig. 1). The CDS problem was initially studied in the setting where the secret is one bit long, and the cost of a CDS scheme is measured by the worst case total amount of communication over all functions $f$, typically as order functions of the input size [2], [5]–[9]. That is, the focus is on the scaling law of the communication complexity as the input size grows to infinity. What is pursued in this work is the traditional Shannon theoretic formulation, where the secret size is allowed to be arbitrarily large, and the communication rate is the number of bits of the secret that can be securely disclosed per bit of total communication. The aim is to characterize the maximum rate, termed the capacity of CDS, for a fixed function $f$.

In [1], we obtain a complete characterization for all functions $f$ where the CDS capacity is the highest, and is equal to $1/2$. In describing this result, we find it convenient to represent the function $f$ by a bipartite graph, where each node denotes a possible signal for certain input and two types (colors) of edges are used to denote whether $f$ is 1 or 0 (see Fig. 1.2). We will use this graph representation of functions $f$ throughout this work. The feasibility condition for capacity $1/2$ is then stated in terms of the graphic properties of $f$. Furthermore, this result is obtained using a novel noise and signal alignment approach, which guides the proof of both (information theoretic) impossibility claims and (linear) protocol designs.

[1]More background on CDS is referred to [1] and references therein.

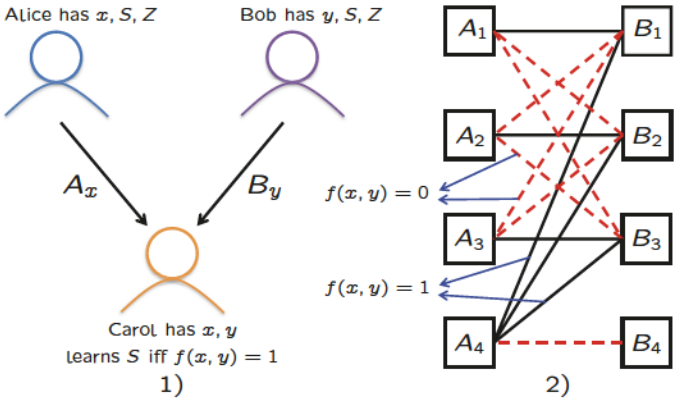


Fig. 1. 1). Alice and Bob (with secret $S$, noise variable $Z$, respective inputs $x, y$) wish to disclose the secret $S$ to Carol if and only if $f(x, y) = 1$ for a binary function $f$, through signals $A_x, B_y$. 2) An example of $f(x, y)$ in graph representation. From pair of nodes connected by a solid black edge (i.e., $f(x, y) = 1$), Carol can decode $S$; from pair of nodes connected by a dashed red edge (i.e., $f(x, y) = 0$), Carol learns nothing about $S$ in the information theoretic sense.

Beyond the best rate scenarios, the simplest uncovered case is also considered in [1] (see Theorem 2), where the linear capacity[2] has been found and this is our starting point. Our goal in this work is to further develop the alignment approach to characterize the linear capacity of a larger class of CDS instances. As our first main result (see Theorem 1), we obtain a general converse bound for linear CDS schemes, which applies to any CDS instance, is parameterized by a covering parameter $\rho$ of the graph representation of $f$, and is equal to $(\rho-1)/(2\rho)$. As our second main result (see Theorem 2), we show that the above converse bound is achievable for a class of graphs, i.e., CDS instances, through a vector linear code based achievable scheme with matching rate. While we find that the converse bound appears to be achievable for more graphs (by verifying a number of examples), an explicit condition of a larger class and a universal code design that applies generally remain elusive. Interestingly, all results are obtained through a more refined view of the alignment approach.

[2]It turns out that the linear capacity, i.e., the highest rate achievable by linear schemes, does not match the best converse bound produced by only Shannon information inequalities [1].

2021 IEEE International Symposium on Information Theory (ISIT) | 978-1-5386-8209-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ISIT45174.2021.9517867

## II. Problem Statement and Preliminaries

Consider a binary function $f(x,y)$, where $(x,y)$ is from some set $\mathcal{I} \subset \{1,2,\cdots,X\} \times \{1,2,\cdots,Y\}$ and its characteristic undirected bipartite graph $G_f = (V,E)$, where the node set $V = \{A_1,\cdots,A_X,\ B_1,\cdots,B_Y\}$ and the edge set $E$ is comprised of the unordered pairs $\{A_x, B_y\}$ such that $(x,y) \in \mathcal{I}$. The edges have two types: if $f(x,y)=1$, $\{A_x,B_y\}$ is a solid black edge and is referred to as a *qualified edge*; if $f(x,y)=0$, $\{A_x,B_y\}$ is a dashed red edge and is referred to as an *unqualified edge* (see Fig. 1.2 for an example).

The variable $x$ ($y$) denotes the input available only to Alice (Bob) and $A_x$ ($B_y$) denotes the signal sent from Alice (Bob) to Carol for securely disclosing the secret $S$, which is comprised of $L$ i.i.d. uniform symbols from a finite field $\mathbb{F}_p$. In addition to the secret $S$, Alice and Bob also hold an independent common noise variable $Z$ (to assist with the secure disclosure task) that is comprised of $L_Z$ i.i.d. uniform symbols from $\mathbb{F}_p$.

$$H(S) = L,\ H(Z) = L_Z,\ \text{(in } p\text{-ary units)} \tag{1}$$
$$H(S,Z) = H(S) + H(Z) = L + L_Z. \tag{2}$$

Each signal $A_x$ ($B_y$) is assumed to be comprised of $N$ symbols from $\mathbb{F}_p$ and must be determined by information available to Alice (Bob).

$$H(A_x, B_y | S, Z) = 0. \tag{3}$$

The disclosure task is said to be successful if the following conditions are satisfied. From a qualified edge, Carol can recover $S$ with no error; from an unqualified edge, Carol must learn nothing about $S$. For all $(x,y) \in \mathcal{I}$, we have

$$[\text{Correctness}]\ \ H(S|A_x,B_y) = 0,\ \text{if } f(x,y)=1; \tag{4}$$
$$[\text{Security}]\ H(S|A_x,B_y) = H(S),\ \text{otherwise } f(x,y)=0. \tag{5}$$

The collection of the mappings from $x,y,S,Z$ to $A_x,B_y$ as specified above is called a CDS scheme.

The CDS rate $R$ characterizes how many symbols of the secret are securely disclosed per symbol of total communication and is defined as follows.

$$R = \frac{L}{2N}. \tag{6}$$

A rate $R$ is said to be achievable if there exists a CDS scheme, for which the correctness and security constraints (4), (5) are satisfied and the rate is no smaller than $R$. The supremum of achievable rates is called the capacity of CDS, $C$.

### A. Graph Definitions

We will use some graphic notions of $G_f = (V,E)$ to state our results, defined as follows. Without loss of generality, we assume that for any node $v \in V$, there exists some node $u \in V$ such that $\{u,v\} \in E$ is an unqualified edge (otherwise, for any $v$ that is connected to only qualified edges, we can set $v$ to be the secret $S$ and then eliminate $v$ and its edges).

*Definition 1 (Qualified/Unqualified Path/Component):* A sequence of distinct connecting qualified (unqualified) edges is called a *qualified (unqualified) path*. A *qualified (unqualified) connected component* is a maximal induced subgraph of $G_f$ such that any two nodes in the subgraph are connected by a qualified (unqualified) path.

*Definition 2 (Internal Qualified Edge):* A qualified edge that connects two nodes in an unqualified path is called an *internal qualified edge*.

For example, in Fig. 1.2, the edge $e = \{A_1, B_1\}$ is an internal qualified edge that connects the two nodes $A_1, B_1$ in the unqualified path $P = \{\{A_1,B_2\}, \{B_2,A_3\}, \{A_3,B_1\}\}$.

*Definition 3 (Connected Edge Cover):* Consider an internal qualified edge $e$ in an unqualified path $P$ and the node set of $P$ is denoted as $V_P \subset V$. A *connected edge cover* of $V_P$ is a set of connected[3] qualified edges $M \subset E$ such that each node in $V_P$ is covered by at least one qualified edge in $M$ and $e \in M$. The size of a connected edge cover for $(e,P)$ is the number of edges in $M$ and is denoted as $\rho(e,P)$. If no such $M$ exists, then $\rho(e,P)$ is defined as $+\infty$. Further, $\rho \triangleq \min_{e,P} \rho(e,P)$.

For example, in Fig. 1.2, consider the internal qualified edge $e = \{A_1,B_1\}$ in the unqualified path $P = \{\{A_1,B_2\}, \{B_2,A_3\}, \{A_3,B_1\}\}$, then the nodes in $P$ are $V_P = \{A_1, B_2, A_3, B_1\}$ and a connected edge cover of $V_P$ is $M = \{\{A_4,B_1\}, \{A_4,B_2\}, \{A_4,B_3\}, \{A_1,B_1\}, \{A_3,B_3\}\}$. In this case, $\rho(e,P) = 5$ as $M$ contains 5 edges and we can verify that the minimum value of $\rho(e,P)$ over all internal qualified edges and unqualified path pairs $(e,P)$ is $\rho = 5$.

### B. Linear Feasibility

The feasibility of a linear CDS scheme is specified below.

**Linear Scheme:** *For a feasible linear CDS scheme, each signal (equivalently, each node $v \in V$)*

$$v = \mathbf{F}_v S + \mathbf{H}_v Z,\ \mathbf{F}_v \in \mathbb{F}_p^{N\times L}, \mathbf{H}_v \in \mathbb{F}_p^{N\times L_Z} \tag{7}$$

*is specified by two precoding matrices, $\mathbf{F}_v$ for the secret $S \in \mathbb{F}_p^{L\times 1}$ and $\mathbf{H}_v$ for[4] the noise $Z \in \mathbb{F}_p^{L_Z\times 1}$ such that the following properties are satisfied.*

- *Consider any edge $\{v,u\}$ and identify the overlap of their noise spaces, i.e., the row space of $\mathbf{H}_v$ and $\mathbf{H}_u$. That is, find matrices $\mathbf{P}_v$ and $\mathbf{P}_u$ such that*

$$\begin{aligned} \mathbf{P}_v\mathbf{H}_v &= \mathbf{P}_u\mathbf{H}_u, \\ \mathit{rank}(\mathbf{P}_v) &= \mathit{dim}(\mathit{rowspan}(\mathbf{H}_v) \cap \mathit{rowspan}(\mathbf{H}_u)), \end{aligned} \tag{8}$$

*then the secret spaces satisfy*

$$[\mathit{Correctness}]\ \ \mathit{rank}(\mathbf{P}_v\mathbf{F}_v - \mathbf{P}_u\mathbf{F}_u) = L, \quad \mathit{if}\ \{u,v\}\ \mathit{is\ qualified}; \tag{9}$$
$$[\mathit{Security}]\ \mathbf{P}_v\mathbf{F}_v = \mathbf{P}_u\mathbf{F}_u,\ \mathit{else}\ \{u,v\}\ \mathit{is\ unqualified}. \tag{10}$$

It is immediate to inspect that the correctness constraint (9) and the security constraint (10) for linear schemes imply the entropic versions (4), (5). Conversely, any feasible linear

[3]That is, any two nodes in $M$ are connected by a qualified path.

[4]Without loss of generality, we assume that $\mathbf{H}_v$ has full row rank, i.e., $\text{rank}(\mathbf{H}_v) = N$, because each $v$ is assumed to connect to at least an unqualified edge so that $I(v;S) = 0$, then the linearly dependent rows of $\mathbf{H}_v$ in $v$ must be linearly dependent as well (thus redundant).

scheme must satisfy (9), (10). Such a linear feasibility framework has appeared in related problems [10], [11]. To facilitate later use, we summarize some useful properties of feasible linear schemes in the following lemma. A detailed proof can be found in Lemma 6 and Lemma 7 of [1].

*Lemma 1:* For any linear scheme as defined above and any edge $\{v, u\}$, we have

$$\text{[Noise Align]} \quad \dim(\text{rowspan}(\mathbf{H}_v) \cap \text{rowspan}(\mathbf{H}_u)) \geq L, \quad \text{if } \{u, v\} \text{ is qualified;} \tag{11}$$

$$\text{[Signal Align]} \quad \mathbf{P}_v \mathbf{F}_v = \mathbf{P}_u \mathbf{F}_u, \text{ if } \{u, v\} \text{ is unqualified.} \tag{12}$$

The intuition of the lemma is as follows. (11) follows from the correctness constraint (9), which requires the overlap of the noise spaces to have at least $L$ dimensions as decoding is only possible over the overlapping space (so referred to as 'noise alignment') and other spaces are covered by independent noise variables. (12) follows from the security constraint (10), which says that over the overlapping noise space, the secret space must also be fully overlapping (so referred to as 'signal alignment' since both noise and secret fully align in this space) as otherwise the unqualified edge can reveal some information of the secret symbols, violating the security constraint.

In the remainder of this paper, we use (9) and (10) to verify the correctness and security of a linear scheme. To illustrate how it works, let us consider again the CDS instance in Fig. 1.2 (reproduced in Fig. 2). We show that rate $R = 2/5$ is achievable, through presenting a vector linear scheme with $L = 4, N = 5$. That is, the secret has $L = 4$ symbols over $\mathbb{F}_3$ ($S = (s_1; s_2; s_3; s_4)$), and each signal has $N = 5$ symb $\mathbb{F}_3$. The assignment of the signals is given in Fig. 2. S

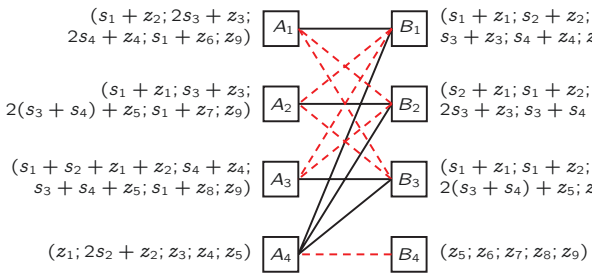


Fig. 2. A CDS instance and the linear scheme of rate $R$ =

Let us verify that the above scheme is correct and For simplicity, we do not write out explicitly the precoding matrices $\mathbf{F}_v$ and $\mathbf{H}_v$ for a signal $v$. Instead, we will directly find the overlap by inspection. Consider qualified edge $\{A_3, B_3\}$. $A_3, B_3$ both contain $(z_1 + z_2; z_4; z_5; z_8)$ (noise overlaps) and can then obtain 4 equations of the secret symbols, $(-s_1 + s_2; s_4; s_3 + s_4; s_1)$, which can recover $S = (s_1; s_2; s_3; s_4)$. Other cases of qualified edges can be verified similarly. Consider the unqualified edge $\{A_3, B_2\}$. $(z_1 + z_2; z_5)$ lies in the overlap of the noise spaces and the secret symbols projecting to this space are both $(s_1 + s_2; s_3 + s_4)$, thus no information is leaked. Other unqualified edges follow similarly. The rate achieved is thus $L/(2N) = 4/10 = 2/5$.

## III. Results

Our first result is a converse bound of linear CDS schemes, parameterized by the minimum connected edge cover number of internal qualified edges, $\rho$ and stated in Theorem 1.

*Theorem 1:* For any CDS instance, the following converse bound holds for all linear schemes.

$$R_{\text{linear}} \leq \frac{\rho - 1}{2\rho}. \tag{13}$$

The proof of Theorem 1 is presented in Section IV.

To give an example, let us consider the CDS instance in Fig. 2. Note that $e = \{A_1, B_1\}$ is an internal qualified edge in unqualified path $P = \{\{A_1, B_2\}, \{B_2, A_3\}, \{A_3, B_1\}\}$, with node set $V_P = \{A_1, B_2, A_3, B_1\}$, where $V_P$ is covered by a connected edge cover $M = \{\{A_4, B_1\}, \{A_4, B_2\}, \{A_4, B_3\}, \{A_1, B_1\}, \{A_3, B_3\}\}$ so that $\rho(e, P) = |M| = 5$ and this edge cover number turns out to be the minimum, i.e., $\rho = 5$. Then Theorem 1 indicates that $R_{\text{linear}} \leq (\rho - 1)/(2\rho) = 2/5$. As rate $2/5$ is linearly achievable (see Fig. 2), the linear capacity of this CDS instance is $2/5$.

Next, we proceed to our second result, which shows that the linear converse in Theorem 1 is tight for a class of CDS instances and is stated in Theorem 2.

*Theorem 2:* For any CDS instance where the qualified edges in each qualified component form either a path or a cycle[5], the linear capacity is $C_{\text{linear}} = (\rho - 1)/(2\rho)$.

Note that Theorem 2 only places constraints on the structure of qualified edges and works for any possible configuration of unqualified edges. The proof is presented in Section V.

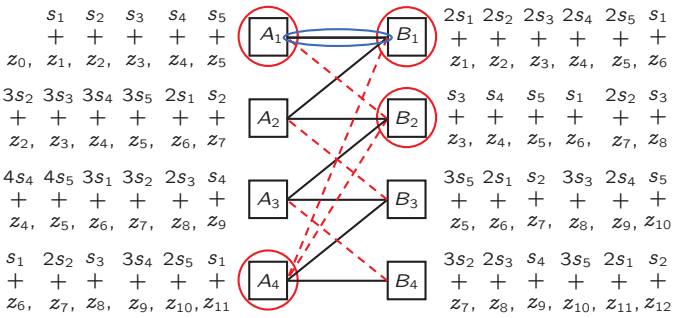


Fig. 3. A CDS instance where the qualified component is a path and a linear capacity achieving scheme.

We give a path example here (and the cycle case is similar) to illustrate the idea. Consider the CDS instance in Fig. 3 and we show that the linear capacity is $C_{\text{linear}} = 5/12$. Theorem 2 can be applied as the instance contains one qualified component, where the qualified edges form a path. $\rho = 6$, because there is an internal qualified edge $e = \{A_1, B_1\}$ (see the blue circle) in unqualified path $P = \{\{A_1, B_2\}, \{B_2, A_4\}, \{A_4, B_1\}\}$ (see the red circles), which is then covered by a qualified path with 6 edges $M = \{\{A_1, B_1\}, \{B_1, A_2\}, \{A_2, B_2\}, \{B_2, A_3\}, \{A_3, B_3\}, \{B_3, A_4\}\}$. It can be verified that this $M$ has the minimum cardinality, so $\rho = 6$. Then the converse bound follows from Theorem 1.

[5]A cycle is a path where the first node is the same as the last node.

We now consider the achievable scheme, where the assignment of each signal is given in Fig. 3. The uniform and i.i.d. noise variables are assigned sequentially to the nodes in the path following a sliding window manner, where the first node $A_1$ uses $z_0, z_1, \cdots, z_5$, the second node uses $z_1, \cdots, z_6$, and so on. The secret symbols $s_1, \cdots, s_5$ are assigned cyclicly to the noise variables, i.e., $(s_1, \cdots, s_5)$ are assigned to $(z_1, \cdots, z_5)$, $(z_6, \cdots, z_{10})$ etc. The coefficients of $s_j$ are the only left and most important part. To this end, focus on each $z_i$ in an arbitrary order and consider only the nodes that contain $z_i$. For example, consider $z_6$, which appears in 6 nodes $B_1, A_2, B_2, A_3, B_3, A_4$ and consider the subgraph induced by these 6 nodes. For the induced subgraph, consider each unqualified component sequentially and assign the same signal to each node in the unqualified component. So here first consider the unqualified path $\{\{B_1, A_4\}, \{A_4, B_2\}\}$ and assign $s_1 + z_6$ to $B_1, A_4, B_2$; second consider the unqualified path $\{A_2, B_3\}$ and assign $2s_1 + z_6$ to $A_2, B_3$; lastly consider $A_3$ (a trivial unqualified component) and assign $3s_1 + z_6$ to $A_3$. All other $z_i$ can be treated in the same manner. This completes the description of the scheme. The security and correctness of the scheme can be inspected from Fig. 3. The rate achieved is $R = L/(2N) = 5/12$ as the secret has $L = 5$ symbols and each signal has $N = 6$ symbols.

## IV. Proof of Theorem 1

For $v_1, \cdots, v_i$, define $\alpha_{v_1 \cdots v_i} \triangleq \dim(\text{rowspan}(\mathbf{H}_{v_1}) \cap \cdots \cap \text{rowspan}(\mathbf{H}_{v_i}))$. Consider any CDS instance $G_f(V, E)$ where $\rho \neq +\infty$ such that there exists an internal qualified edge $e$ in an unqualified path $P$ and $\rho(e, P) = \rho$. Then the connected edge cover $M$ for nodes $V_P$ in $P$ contains $\rho$ edges and $\rho + 1$ nodes, denoted as $V_M = \{v_1, v_2, \cdots, v_{\rho+1}\}$ Note that $M$ has minimal size and is a spanning tree of nodes $V_M$.

Start with the internal qualified edge $e$ in $M$, say $e = \{v_{i_1}, v_{i_2}\} \subset M, i_1, i_2 \in \{1, 2, \cdots, \rho+1\}$. As $M$ is connected, there must exist a node $v_{i_3} \in V_M, i_3 \notin \{i_1, i_2\}$ and a node $u_1 \in \{v_{i_1}, v_{i_2}\}$ such that $\{u_1, v_{i_3}\}$ is a qualified edge. Then from sub-modularity, we have

$$\alpha_{v_{i_1} v_{i_2} v_{i_3}} \geq \alpha_{v_{i_1} v_{i_2}} + \alpha_{u_1 v_{i_3}} - N. \tag{14}$$

Then we proceed similarly to find $v_{i_4} \in V_M, i_4 \notin \{i_1, i_2, i_3\}$ such that $\{u_2, v_{i_4}\}$ is a qualified edge, where $u_2 \in \{v_{i_1}, v_{i_2}, v_{i_3}\}$. Again from sub-modularity, we have

$$\begin{aligned} \alpha_{v_{i_1} v_{i_2} v_{i_3} v_{i_4}} &\geq \alpha_{v_{i_1} v_{i_2} v_{i_3}} + \alpha_{u_2 v_{i_4}} - N & (15) \\ &\overset{(14)}{\geq} \alpha_{v_{i_1} v_{i_2}} + \alpha_{u_1 v_{i_3}} + \alpha_{u_2 v_{i_4}} - 2N. & (16) \end{aligned}$$

Continue this procedure, i.e., we include one node $v_{i_j} \in V_M, i_j \notin \{i_1, \cdots, i_{j-1}\}, j \in \{5, \cdots, \rho+1\}$ at one time such that $\{u_{j-2}, v_{i_j}\} \in M$ and $u_{j-2} \in \{v_{i_1}, \cdots, v_{i_{j-1}}\}$. Then

$$\begin{aligned} \alpha_{v_{i_1} v_{i_2} \cdots v_{i_{\rho+1}}} &\geq \alpha_{v_{i_1} \cdots v_{i_\rho}} + \alpha_{u_{\rho-1} v_{i_{\rho+1}}} - N \geq \cdots \\ &\geq \alpha_{v_{i_1} v_{i_2}} + \alpha_{u_1 v_{i_3}} + \alpha_{u_2 v_{i_4}} + \cdots \\ &\quad + \alpha_{u_{\rho-1} v_{i_{\rho+1}}} - (\rho - 1)N. \end{aligned} \tag{17}$$

Note that $i_1, \cdots, i_{\rho+1}$ are distinct so that $V_M = \{v_1, \cdots, v_{\rho+1}\} = \{v_{i_1}, \cdots, v_{i_{\rho+1}}\}$. As the $\rho+1$ noise spaces have an overlap of dimension $\alpha_{v_{i_1} v_{i_2} \cdots v_{i_{\rho+1}}}$, there exist $\rho + 1$ projection matrices $\mathbf{P}^{\cap}_{v_{i_1}}, \cdots, \mathbf{P}^{\cap}_{v_{i_{\rho+1}}}$ of rank $\alpha_{v_{i_1} v_{i_2} \cdots v_{i_{\rho+1}}}$ each such that

$$\mathbf{P}^{\cap}_{v_{i_1}} \mathbf{H}_{v_{i_1}} = \mathbf{P}^{\cap}_{v_{i_2}} \mathbf{H}_{v_{i_2}} = \cdots = \mathbf{P}^{\cap}_{v_{i_{\rho+1}}} \mathbf{H}_{v_{i_{\rho+1}}}. \tag{18}$$

Next, switch focus to the unqualified path $P$. Consider the nodes $V_P \subset V_M$ and denote $V_P = \{v_{i_1}, v_{i_2}, v_{j_1}, v_{j_2}, \cdots, v_{j_{|V_P|-2}}\} \subset \{v_{i_1}, v_{i_2}, \cdots, v_{i_{\rho+1}}\} = V_M$. By (12), i.e., the signal alignment constraint from Lemma 1, and (18),

$$\mathbf{P}^{\cap}_{v_{i_1}} \mathbf{F}_{v_{i_1}} = \mathbf{P}^{\cap}_{v_{i_2}} \mathbf{F}_{v_{i_2}}. \tag{19}$$

Finally, consider the internal qualified edge $e = \{v_{i_1}, v_{i_2}\}$ and identify the noise overlap through matrices $\mathbf{P}_{v_{i_1}}, \mathbf{P}_{v_{i_2}}$ that have rank $\alpha_{v_{i_1}, v_{i_2}}$, i.e., $\mathbf{P}_{v_{i_1}} \mathbf{H}_{v_{i_1}} = \mathbf{P}_{v_{i_2}} \mathbf{H}_{v_{i_2}}$. Noting that rowspan($\mathbf{P}^{\cap}_{v_{i_1}}$) is a subspace of rowspan($\mathbf{P}_{v_{i_1}}$), we set

$$\begin{aligned} \mathbf{P}^{\cap}_{v_{i_1}} &= \mathbf{P}_{v_{i_1}}(1 : \alpha_{v_{i_1} v_{i_2} \cdots v_{i_{\rho+1}}}, :), \\ \mathbf{P}^{\cap}_{v_{i_2}} &= \mathbf{P}_{v_{i_2}}(1 : \alpha_{v_{i_1} v_{i_2} \cdots v_{i_{\rho+1}}}, :) \end{aligned} \tag{20}$$

without loss of generality, i.e., the first $\alpha_{v_{i_1} v_{i_2} \cdots v_{i_{\rho+1}}}$ rows of $\mathbf{P}_{v_{i_1}}$ are $\mathbf{P}^{\cap}_{v_{i_1}}$. Then from the correctness constraint (9) for qualified edge $e = \{v_{i_1}, v_{i_2}\}$, we have

$$\begin{aligned} L &\overset{(9)}{=} \text{rank}\left(\mathbf{P}_{v_{i_1}} \mathbf{F}_{v_{i_1}} - \mathbf{P}_{v_{i_2}} \mathbf{F}_{v_{i_2}}\right) \\ &\overset{(19)(20)}{=} \text{rank}\Big(\mathbf{P}_{v_{i_1}}(\alpha_{v_{i_1} v_{i_2} \cdots v_{i_{\rho+1}}} + 1 : \alpha_{v_1 v_2}, :)\mathbf{F}_{v_{i_1}} \\ &\qquad - \mathbf{P}_{v_{i_2}}(\alpha_{v_{i_1} v_{i_2} \cdots v_{i_{\rho+1}}} + 1 : \alpha_{v_1 v_2}, :)\mathbf{F}_{v_{i_2}}\Big) \\ &\leq \alpha_{v_{i_1} v_{i_2}} - \alpha_{v_{i_1} v_{i_2} \cdots v_{i_{\rho+1}}} \\ &\overset{(17)}{\leq} \alpha_{v_{i_1} v_{i_2}} - \Big(\alpha_{v_{i_1} v_{i_2}} + \alpha_{u_1 v_{i_3}} + \alpha_{u_2 v_{i_4}} + \cdots \\ &\qquad + \alpha_{u_{\rho-1} v_{i_{\rho+1}}} - (\rho - 1)N\Big) \\ &= (\rho - 1)N - \Big(\alpha_{u_1 v_{i_3}} + \cdots + \alpha_{u_{\rho-1} v_{i_{\rho+1}}}\Big) \\ &\overset{(11)}{\leq} (\rho - 1)N - (\rho - 1)L \\ &\Rightarrow \quad R_{\text{linear}} = L/(2N) \leq (\rho - 1)/(2\rho). \end{aligned}$$

## V. Proof of Theorem 2

We present a vector linear CDS scheme that achieves rate $(\rho - 1)/(2\rho)$ whe each qualified component of the CDS instance is a path or cycle. Specifically, we set $L = \rho - 1$, i.e., each secret has $L$ symbols $S = (s_1, \cdots, s_{\rho-1})$ from $\mathbb{F}_p$ and $N = \rho$, i.e., each signal (node) $v$ has $N$ symbols from $\mathbb{F}_p$. We assume that $p$ is a prime number and $p \geq 2\rho - 2$.

Define $l_1, \cdots, l_{\rho-1}$ as $L$ generic linear combinations of $S$.

$$\begin{aligned} (l_1; \cdots; l_{\rho-1}) &= \mathbf{C}_{(\rho-1)\times(\rho-1)} \times (s_1; \cdots, s_{\rho-1}) \\ \mathbf{C}_{(\rho-1)\times(\rho-1)}(i, j) &= \frac{1}{x_i - y_j}, i, j \in \{1, \cdots, \rho - 1\} \end{aligned} \tag{21}$$

where $x_i, y_j$ are distinct elements from $\mathbb{F}_p$, so $\mathbf{C}_{(\rho-1)\times(\rho-1)}$ is a Cauchy matrix.

Consider any CDS instance $G_f(V, E)$ such that the minimum connected edge cover number for any internal qualified edge is $\rho$. Suppose the instance contains $Q$ qualified

components, where each qualified component is either a path or a cycle of qualified edges. Denote the node set of the $q$-th qualified component by $V^q, q \in \{1, \cdots, Q\}$ such that $V = V^1 \cup \cdots \cup V^Q$. For each qualified component, we will use independent uniform i.i.d. noise symbols from $\mathbb{F}_p$, denoted as $z^q = (z_0^q, z_1^q, z_2^q, \cdots)$. So $Z = (z^1, \cdots, z^Q)$. We are now ready to specify the signal design.

1. Consider each qualified component sequentially. If the $q$-th qualified component is a path, go to 2; otherwise the $q$-th qualified component is a cycle, go to 3.

2. The nodes $V^q = \{v_1^q, \cdots, v_{|V^q|}^q\}$ form a path. Suppose $\{v_1^q, v_2^q\}, \{v_2^q, v_3^q\}, \cdots, \{v_{|V^q|-1}^q, v_{|V^q|}^q\}$ are qualified edges.

2.1. Assign the noise variables sequentially as follows.

$$v_1^q = (z_0^q, z_1^q, \cdots, z_{\rho-1}^q), v_2^q = (z_1^q, z_2^q, \cdots, z_\rho^q), \\ \cdots, v_{|V^q|}^q = (z_{|V^q|-1}^q, \cdots, z_{|V^q|+\rho-2}^q). \quad (22)$$

2.2. We now include the secret symbols to each node. Consider the nodes that contain each noise symbol $z_1^q$, $\cdots$, $z_{|V^q|+\rho-2}^q$ sequentially and the induced subgraph formed by these nodes. Note that each noise symbol $z_j^q$, $j \in \{1, \cdots, |V^q| + \rho - 2\}$ appears at no more than $\rho$ nodes and denote the induced subgraph by $G_j^q \subset G_f$. Suppose $G_j^q$ contains $K_j^q$ unqualified components[6]. For each node $v_i^q$ in the $k$-th unqualified component of $G_j^q$, $k \in \{1, \cdots, K_j^q\}$, $j \in \{1, \cdots, |V^q| + \rho - 2\}$, replace $z_j^q$ by $k \times s_{j \bmod (\rho-1)} + z_j^q$. Note that in $s_{j \bmod (\rho-1)}$, the subscript is defined over $\{1, \cdots, \rho - 1\}$ as the secret symbols are $S = (s_1, \cdots, s_{\rho-1})$.

3. The nodes $V^q = \{v_1^q, \cdots, v_{|V^q|}^q\}$ form a cycle. Suppose $\{v_1^q, v_2^q\}, \cdots, \{v_{|V^q|-1}^q, v_{|V^q|}^q\}, \{v_{|V^q|}^q, v_1^q\}$ are qualified edges.

3.1. Assign the noise variables cyclicly as follows.

$$v_1^q = (z_1^q, z_2^q, \cdots, z_\rho^q), v_2^q = (z_2^q, z_3^q, \cdots, z_{\rho+1}^q), \\ \cdots, v_{|V^q|}^q = (z_{|V_q|}^q, z_1^q, \cdots, z_{\rho-1}^q). \quad (23)$$

3.2. We now include the secret symbols to each node. Consider the nodes that contain each noise symbol $z_1^q, \cdots, z_{|V^q|}^q$ sequentially and the induced subgraph formed by these nodes. Note that each noise symbol $z_j^q, j \in \{1, \cdots, |V^q|\}$ appears at $\rho$ nodes and denote the induced subgraph by $G_j^q \subset G_f$. Suppose $G_j^q$ contains $K_j^q$ unqualified components. For each node $v_i^q$ in the $k$-th unqualified component of $G_j^q$, $k \in \{1, \cdots, K_j^q\}$, if $j \in \{1, \cdots, \rho-1\}$, replace $z_j^q$ by $k \times l_j + z_j^q$; otherwise $j \in \{\rho, \cdots, |V_q|\}$, replace $z_j^q$ by $k \times s_{j \bmod (\rho-1)} + z_j^q$.

After describing the signal design, we proceed to show that the scheme is correct and secure and complete the proof.

First, we prove that the correctness constraint (9) is satisfied. We consider each qualified component and have two cases.

1. The qualified component is a path. From the noise assignment (22), we know that the two nodes $u, v$ in any qualified edge share $L = \rho - 1$ noise symbols with consecutive subscripts. Further, according to the signal assignment, these $L$ consecutive noise symbols are each mixed with one distinct secret symbol from the $L$ symbols in $S$. In addition, each shared secret symbol $s_i, i \in \{1, \cdots, L\}$ in $v$ and $u$ is multiplied by different coefficients $k$. We prove this claim by contradiction, i.e., suppose that the coefficients $k$ are the same. Then $e = \{u, v\}$ must be an internal qualified edge in an unqualified path $P$, and we can find a connected edge cover $M$ for the nodes in $P$ and all nodes in $M$ share one same noise symbol. Recall from Definition 3 that $M$ contains at least $\rho + 1$ nodes while these nodes share one same noise symbol, which is not possible because from the noise assignment (22), each noise symbol only appears at $\rho$ nodes at most. Thus the coefficients for the $L$ secret symbols in $v, u$ are all distinct and from $\{v, u\}$ we can recover $S$ with no error.

2. The second case is when the qualified component is a cycle, whose proof is similar to the path case. Similarly from the noise assignment (23), any two nodes $u, v$ in a qualified edge share $L = \rho - 1$ noise symbols with cyclicly consecutive subscripts. Further, according to the signal assignment, these $L$ noise symbols are each mixed with either one distinct secret symbol $s_i$ from the $L$ symbols in $S$ or one generic linear combination $l_j$. With a similar reasoning as above (due to the definition of $\rho$ and each noise appears at $\rho$ nodes), the multiplicative coefficients $k$ for $s_i, l_j$ are distinct. As $l_j$ are from a Cauchy matrix (see (21)), whose every square sub-matrix has full rank [12], from $\{v, u\}$ we can obtain $L$ equations of form $s_i, l_j$ thus recover $S$ with no error.

Second, we prove that the security constraint (10) is satisfied. We have two cases for an unqualified edge.

1. The two nodes $u, v$ of the unqualified edge are from the same qualified component. Security is guaranteed because in the signal assignment, when the noise space overlaps, the same signal equation is assigned, i.e., signal alignment is ensured and (10) holds.

2. The two nodes $u, v$ of the unqualified edge are from two different qualified components. As the noise symbols $z^q, z^{q'}$ are independent for distinct qualified components, the noise spaces of $u, v$ have no overlap and (10) trivially holds.

## VI. Discussion

We take a Shannon theoretic perspective at the canonical CDS problem to seek capacity characterizations where the secret size is allowed to approach infinity. This Shannon theoretic perspective follows the footsteps of recent attempts in information theory on other cryptographic primitives [13]–[22]. To this end, we further develop the noise and signal alignment approach (introduced in [1]), which is a variation of interference alignment originally studied in wireless communication [23]–[25], to characterize the linear capacity of a class of CDS instances, which go beyond the highest capacity scenarios found in [1]. Along the line, we identify a general linear converse bound (see Theorem 1) and a linear feasibility framework that facilitates the design of linear schemes once the target rate value is fixed (see Section II-B). However, these results are not sufficient to fully understand the linear capacity of CDS in general - an open problem.


## Acknowledgement

This work is supported in part by NSF grant CCF-2007108.


[6]A node that connects to no unqualified edge is a trivial unqualified component. As there are at most $\rho$ nodes in $G_j^q$, we have that $K_j^q \leq \rho$.

## REFERENCES


[1] Z. Li and H. Sun, "Conditional Disclosure of Secrets: A Noise and Signal Alignment Approach," *arXiv preprint arXiv:2002.05691*, 2020.

[2] Y. Gertner, Y. Ishai, E. Kushilevitz, and T. Malkin, "Protecting Data Privacy in Private Information Retrieval Schemes," in *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*. ACM, 1998, pp. 151–160.

[3] T. Liu, V. Vaikuntanathan, and H. Wee, "Towards Breaking the Exponential Barrier for General Secret Sharing," in *Annual International Conference on the Theory and Applications of Cryptographic Techniques*. Springer, 2018, pp. 567–596.

[4] B. Applebaum, A. Beimel, O. Nir, and N. Peter, "Better Secret Sharing via Robust Conditional Disclosure of Secrets," in *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, 2020, pp. 280–293.

[5] R. Gay, I. Kerenidis, and H. Wee, "Communication Complexity of Conditional Disclosure of Secrets and Attribute-Based Encryption," in *Annual Cryptology Conference*. Springer, 2015, pp. 485–502.

[6] B. Applebaum, B. Arkis, P. Raykov, and P. N. Vasudevan, "Conditional Disclosure of Secrets: Amplification, Closure, Amortization, Lower-Bounds, and Separations," in *Annual International Cryptology Conference*. Springer, 2017, pp. 727–757.

[7] S. Laur and H. Lipmaa, "A New Protocol for Conditional Disclosure of Secrets and Its Applications," in *International Conference on Applied Cryptography and Network Security*. Springer, 2007, pp. 207–225.

[8] B. Applebaum and P. N. Vasudevan, "Placing Conditional Disclosure of Secrets in the Communication Complexity Universe," in *10th Innovations in Theoretical Computer Science Conference (ITCS 2019)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018.

[9] T. Liu, V. Vaikuntanathan, and H. Wee, "Conditional Disclosure of Secrets via Non-Linear Reconstruction," in *Annual International Cryptology Conference*. Springer, 2017, pp. 758–790.

[10] S. H. Dau, V. Skachek, and Y. M. Chee, "On the Security of Index Coding with Side Information," *IEEE Transactions on Information Theory*, vol. 58, no. 6, pp. 3975–3988, 2012.

[11] H. Sun, "Secure Groupcast: Extra-Entropic Structure and Linear Feasibility," *arXiv preprint arXiv:2006.05944*, 2020.

[12] S. Schechter, "On the inversion of certain matrices," *Mathematical Tables and Other Aids to Computation*, vol. 13, no. 66, pp. 73–77, 1959.

[13] 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.

[14] E. J. Lee and E. Abbe, "Two Shannon-Type Problems on Secure Multiparty Computations," in *2014 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, pp. 1287–1293.

[15] D. Data, V. M. Prabhakaran, and M. M. Prabhakaran, "Communication and Randomness Lower Bounds for Secure Computation," *IEEE Transactions on Information Theory*, vol. 62, no. 7, pp. 3901–3929, 2016.

[16] Y. Zhou, H. Sun, and S. Fu, "On the Randomness Cost of Linear Secure Computation," in *2019 53rd Annual Conference on Information Sciences and Systems (CISS)*, March 2019, pp. 1–6.

[17] Y. Zhao and H. Sun, "Expand-and-Randomize: An Algebraic Approach to Secure Computation," *arXiv preprint arXiv:2001.00539*, 2020.

[18] H. Sun, "The Capacity of Anonymous Communications," *IEEE Transactions on Information Theory*, vol. 65, no. 6, pp. 3871–3879, 2018.

[19] B. Tahmasebi and M. A. Maddah-Ali, "Private Sequential Function Computation," *arXiv preprint arXiv:1908.01204*, 2019.

[20] Z. Wang, K. Banawan, and S. Ulukus, "Private Set Intersection: A Multi-Message Symmetric Private Information Retrieval Perspective," *arXiv preprint arXiv:1912.13501*, 2020.

[21] ——, "Multi-party private set intersection: An information-theoretic approach," *arXiv preprint arXiv:2008.07504*, 2020.

[22] H. Sun, "Compound Secure Groupcast: Key Assignment for Selected Broadcasting," *arXiv preprint arXiv:2004.14986*, 2020.

[23] S. A. Jafar, "Interference Alignment - A New Look at Signal Dimensions in a Communication Network," *Foundations and Trends in Communications and Information Theory*, vol. 7, no. 1, pp. 1–134, 2011. [Online]. Available: http://dx.doi.org/10.1561/0100000047

[24] ——, "Topological Interference Management through Index Coding," *IEEE Trans. on Inf. Theory*, vol. 60, no. 1, pp. "529–568", Jan. 2014.

[25] H. Sun and S. A. Jafar, "Index Coding Capacity: How far can one go with only Shannon Inequalities?" *IEEE Trans. on Inf. Theory*, vol. 61, no. 6, pp. 3041–3055, 2015.