

Byzantine-Eavesdropper Alliance: How to Achieve Symmetric Privacy in Quantum X -Secure B -Byzantine E -Eavesdropped U -Unresponsive T -Colluding PIR?

Mohamed Nomeir^{ID}, Graduate Student Member, IEEE, Alptug Aytekin^{ID}, Graduate Student Member, IEEE, and Sennur Ulukus^{ID}, Fellow, IEEE

Abstract—We consider the quantum *symmetric* private information retrieval (QSPIR) problem in a system with N databases and K messages, with U unresponsive servers, T -colluding servers, and X -security parameter, under several fundamental threat models. In the first model, there are $\mathcal{E}_1$ eavesdropped links in the uplink direction (the direction from the user to the N servers), $\mathcal{E}_2$ eavesdropped links in the downlink direction (the direction from the servers to the user), where $|\mathcal{E}_1|, |\mathcal{E}_2| \leq E$; we coin this eavesdropper setting as *dynamic* eavesdroppers. We show that super-dense coding gain can be achieved for some regimes. In the second model, we consider the case with Byzantine servers, i.e., servers that can coordinate to devise a plan to harm the privacy and security of the system together with static eavesdroppers, by listening to the same links in both uplink and downlink directions. It is important to note the considerable difference between the two threat models, since the eavesdroppers can take huge advantage of the presence of the Byzantine servers. Unlike the previous works in SPIR with Byzantine servers, that assume that the Byzantine servers can send only random symbols independent of the stored messages, we follow the original definition of Byzantine servers in PIR proposed by Banawan-Ulukus, where the Byzantine servers can send symbols that can be functions of the storage, queries, as well as the random symbols in a way that can produce worse harm to the system. In the third and the most novel threat model, we consider the presence of Byzantine servers and dynamic eavesdroppers together. We show that having dynamic eavesdroppers along with Byzantine servers in the same system model creates more threats to the system than having static eavesdroppers with Byzantine servers. This is the first work that considers the quantum version of unresponsive and eavesdropped threat model. In addition, this is the first work, classical or quantum, that considers the presence of Byzantine servers together with eavesdroppers in the same system model (static or dynamic eavesdroppers). Another layer of difficulty that we handle in this work stems from the *symmetric* privacy requirement in the presence of Byzantine servers, which by itself has never been studied before in classical or quantum variations;

here the Byzantine servers may attempt to leak information about undesired messages to the user, which is not allowed in SPIR.

Index Terms—Quantum information theory, private information retrieval, security, eavesdropping.

I. INTRODUCTION

A. Problem Introduction and Our Contributions

RECENTLY, the quantum private information retrieval (QPIR) problem was introduced [2]. In that setting, there are N servers that share an entangled state. A user, who wishes to retrieve a message privately from K messages, sends queries over classical channels to each server independently. Upon receiving the queries, the servers, who share a globally known entangled state, apply quantum operations on their individual quantum subsystems and send them to the user over separate quantum channels. It was shown in [2] that the capacity for the quantum variation with symmetric privacy (QSPIR) is 1 when $N = 2$. In [3], the T -colluding QPIR is investigated and it was shown that the rate is $R = \min \{1, 2 \left(\frac{N-T}{N} \right)\}$. More recently, a new mathematical abstraction has been developed to model the system with N entangled transmitters in a classical quantum setting that transmits over N noise-free separate quantum channels to one receiver where the encoding of the classical information is done using unitary Pauli operations and the decoding is done using their corresponding projective value measurements (PVMs) [4]. This abstraction is coined as *N -sum box abstraction*. It was shown in [4] that the quantum variation of the X -secure, T -colluding PIR (QXTPIR) achieves a rate equal to $R_Q = \min \{1, 2 \left(\frac{N-X-T}{N} \right)\}$.

This paper first investigates how to merge the dynamic eavesdropper threat with unresponsive servers in a single framework. Afterward, a new model that is not studied in classical or quantum symmetric PIR (SPIR) settings is investigated. In particular, we analyze the interactions between Byzantine servers, as defined in [1], and eavesdroppers in a QSPIR system. We consider two eavesdropper models. In the first kind, the eavesdropper is *static*, where the uplinks and downlinks listened to are identical [5], [6]. In the second kind, the eavesdropper is *dynamic*, where the eavesdropped links can be different for the uplink and downlink direc-

Received 8 December 2024; revised 9 September 2025; accepted 15 September 2025. Date of publication 17 October 2025; date of current version 21 May 2026. An earlier version of this paper was presented in part at the 2025 IEEE International Conference on Communications [DOI: 10.1109/ICC52391.2025.11161059]. (Corresponding author: Sennur Ulukus.)

The authors are with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: mnomeir@umd.edu; aaytekin@umd.edu; ulukus@umd.edu).

Communicated by M. Skoglund, Associate Editor for Security and Privacy. Digital Object Identifier 10.1109/TIT.2025.3623081

tions. The latter is considered only in [7]. We find the rate regions for both cases for the general case of the quantum U -unresponsive, T -colluding, and X -secure SPIR model. This is the first work that considers this kind of Byzantine behavior under symmetric privacy constraints in a classical or quantum setting. In addition, this is the first work that investigates this kind of Byzantine server with static or dynamic eavesdroppers.

B. Related Work

Private information retrieval (PIR) is the problem that considers the retrieval of a message by a user, out of K messages stored in a replicated manner in N servers, without revealing the required message index to any individual server [8]. The main performance metric for PIR is the ratio between the number of required message symbols and the total number of symbols downloaded, coined as the PIR rate. In [9], it was shown that the optimal rate, i.e., the capacity, of PIR is $(1 + \frac{1}{N} + \frac{1}{N^2} + \dots + \frac{1}{N^{K-1}})^{-1}$. In [10], another requirement was added, coined symmetric privacy. In symmetric PIR (SPIR), the servers want the other, non-required, messages to be hidden from the user. It was shown that the capacity of SPIR is $1 - \frac{1}{N}$ [10]. Another variation was formulated in [11], coined T -colluding PIR (TPIR), where any T out of N servers can collude to share their queries received from the user to break the user privacy. It was shown that the capacity of TPIR is $(1 + \frac{T}{N} + \frac{T^2}{N^2} + \dots + \frac{T^{K-1}}{N^{K-1}})^{-1}$. In [12], the general problem of TSPIR was solved and the capacity is given by $1 - \frac{T}{N}$. Another version for collusion was defined in [13] for arbitrary collusion patterns. In [1], the Byzantine version of TPIR problem (BTPIR) was formulated, where any B out of N servers can act as Byzantine servers and respond to the user queries with arbitrarily errored answers. It was shown that the capacity of BTPIR is $(N - 2B)^{K-1} (1 + \frac{T}{N-2B} + \dots + \frac{T^{K-1}}{N-2B^{K-1}})^{-1}$. In [14], the eavesdropped version of TSPIR was defined including active eavesdroppers. It was shown that if there are E passive eavesdropped links and A active eavesdropped links, then the capacity for ETSPPIR is $1 - \frac{2A + \max(E, T)}{N}$. Subsequently, the capacity of the ETPIR, i.e., with no symmetry requirement, was found in [5]. Another interesting model for eavesdroppers can be found in [6], where a wiretap channel viewpoint is used to model the eavesdroppers.

It is important to note that there is a significant difference between the servers being Byzantine and the eavesdroppers being active. In the simplest sense, the Byzantine servers can help the eavesdropper get some information about the messages and the queries. However, an active eavesdropper can act on the answers being transmitted by the servers to cause errors at the user side. In addition, in the symmetric privacy setting, there is a shared randomness among all servers, which causes more issues when the servers themselves are Byzantine. Another variant was studied in [15], where any U servers can be unresponsive, and B can be Byzantine (UBTPIR). The achievable rate is given by $1 - \frac{N-U-T-2B}{N}$, and is conjectured to remain the same in the symmetric privacy setting. The main difference between that work and our work,

aside from the presence of an eavesdropper, is that in [15] the Byzantine servers send random symbols independently of the other servers, the storage, and the queries. It is clear that our model here is more strict since we assume possible cooperation between the Byzantine servers. In the collaborative Byzantine servers setting, if any Byzantine server decides to send a message symbol randomly, the symmetric privacy requirement is violated. The latest model of PIR is X -secure TPIR (XTPIR) in [16]. The exact capacity of this setting is still an open problem. The highest achievable rate to date is $1 - \frac{X+T}{N}$ [17]. Generally, for the case of X -secure, B -Byzantine, U -unresponsive, T -colluding PIR (XBUTPIR), the state of the art rate is given by [18]. More variations and results can be found in [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], and [32], in addition to the newly developed approaches using algebraic geometric codes in PIR in [33] and [34].

As for the QPIR literature, in [7] and [35] the eavesdropper variation was studied and the rate regions were found. In [36], the quantum unresponsive servers variation was analyzed and the rate regions were given. Finally, the quantum Byzantine servers version was analyzed and the rate regions were found in [37]. More applications for the new N -sum box abstraction can be found in [38], [39], and [40].

II. PROBLEM FORMULATION

We consider a system consisting of N servers and K messages. The messages are stored classically within the servers. Each message W_i consists of L symbols chosen uniformly and independently at random from the field $\mathbb{F}_q$, i.e.,

$$H(W_i) = L, \quad (1)$$

$$H(W_{[K]}) = KL, \quad (2)$$

where $H(\cdot)$ is the Shannon entropy with base q [41].

Let S_n denote the storage at the n th server. Any $\mathcal{X} \subset [N]$ servers can communicate to attempt to decode the stored content where $|\mathcal{X}| \leq X$. Thus, the message content must be hidden from any X communicating servers, i.e.,

$$I(W_{[K]}; S_{\mathcal{X}}) = 0. \quad (3)$$

A user chooses θ uniformly at random from $[K]$ to retrieve the θ th message. The user sends the query $Q_n^{[\theta]}$ over a classical channel to the n th server. Any $\mathcal{T} \subset [N]$ servers can collude to attempt to decode the required message index θ , where $|\mathcal{T}| \leq T$. Thus, the queries must be secure against any $\mathcal{T}$ colluding servers, i.e.,

$$I(Q_{\mathcal{T}}; \theta) = 0. \quad (4)$$

The servers share a quantum system $\mathcal{A}^{[0]} = \mathcal{A}_1^{[0]} \dots \mathcal{A}_N^{[0]}$ represented by the quantum state ρ prior to the initiation of the scheme. The honest but curious servers encode their answers over their own part of the state $\rho_n = \text{tr}_{i \in [N] \setminus n} \rho$ and reply back over their separate quantum channels to the user using their individual stored symbols S_n , their received queries $Q_n^{[\theta]}$, and the shared common randomness among all the servers Z' , i.e., the honest servers generate the classical answers such that

$$H(A_n | S_n, Q_n, Z') = 0, \quad (5)$$

and encode them over their separate quantum systems, i.e.,

$$\mathcal{A}_n = \text{Enc}_n(A_n, \mathcal{A}_n^{[0]}), \quad (6)$$

where $\text{Enc}_n(\cdot, \cdot)$ is the quantum encoder at the n th server.

In addition, the Byzantine servers generate classical answers as an arbitrary function of their received queries, storage, shared common randomness, and an arbitrary random variable Δ_n , i.e.,

$$H(A_n|S_n, Q_n, Z', \Delta_n) = 0, \quad (7)$$

and

$$\mathcal{A}_n = \text{Enc}'_n(A_n, \mathcal{A}_n^{[0]}), \quad (8)$$

where $\text{Enc}'_n(\cdot, \cdot)$ is any arbitrary quantum encoder.

Remark 1: As mentioned earlier, here we consider Byzantine servers that send answers not just based on an independent random variable but can be any arbitrary function of the storage, queries, masking random variables, and an arbitrary random variable.

Upon receiving all the answers, the user must be able to decode the required message W_θ after applying the quantum measurement operations on the received quantum answers, i.e.,

$$H(W_\theta|Y, Q_{[N]}) = 0, \quad (9)$$

where Y is the result of the measurements on the received quantum answers.

In addition, to achieve symmetric privacy, the user cannot get any information about messages other than the desired message W_θ , i.e.,

$$S(W_{[K] \setminus \{\theta\}}; \mathcal{A}_{[N]}, Q_{[N]}) = 0. \quad (10)$$

Note that (10) includes both honest and Byzantine servers.

Furthermore, there are E eavesdropped uplinks and E eavesdropped downlinks where the passive eavesdropper can listen to any set of queries and answers. Both the index and the messages should be secure against the eavesdropper, i.e.,

$$S(W_{[K]}, \theta; Q_{\mathcal{E}_1}, \mathcal{A}_{\mathcal{E}_2}) = 0, \quad (11)$$

where $\mathcal{E}_1, \mathcal{E}_2 \subset [N]$, and $|\mathcal{E}_1|, |\mathcal{E}_2| \leq E$.

Prior to the presentation of our main results in the next section, we make the following important remarks.

Remark 2: As stated earlier, in the PIR and QPIR literature, two different kinds of eavesdropper models have been considered. In the first kind, the eavesdropper is *static*, which means that the uplinks and downlinks listened to are identical, i.e., $\mathcal{E}_1 = \mathcal{E}_2$ in (11). In the second kind, the eavesdropper is *dynamic*, which means that the eavesdropped links can be different for the uplink and the downlink. The latter is considered only in [7] so far in the literature. However, as analyzed in [7], the static or dynamic eavesdroppers are essentially the same when acting alone with no other harmful entity.

Remark 3: If there are no eavesdroppers in the system, the Byzantine servers have fewer options for actions to harm the system, since the scheme needs only to protect for the user and the databases, i.e., symmetric, privacy. However, when there is a dynamic eavesdropper in the system, the Byzantine servers can also relay part of the queries they have received.

As a simple example, consider the case that the eavesdropper listens to the servers $\mathcal{E}_1$ in the uplink and the servers $\mathcal{E}_2$ in the downlink, where $\mathcal{B} \subset \mathcal{E}_2$, and $\mathcal{E}_1 \cap \mathcal{E}_2 = \emptyset$, with $|\mathcal{E}_1|, |\mathcal{E}_2| \leq E$. In this case, if the queries are protected against E links only and if the Byzantine servers decide to re-transmit part of the queries, then the scheme will be compromised.

III. MAIN RESULTS

Theorem 1: For a QSPIR system with N databases, where X of them can communicate to decode messages, $|\mathcal{E}_1| \leq E$ of the uplinks and $|\mathcal{E}_2| \leq E$ of the downlinks are tapped by a passive eavesdropper, U of the databases are unresponsive, and T of them are colluding to identify user-required message index, the following rate is achievable with symmetric privacy,

$$R_Q = \begin{cases} 2 \left(\frac{N-X-M-U}{N} \right), & N-U > X+M \geq \frac{N}{2}, \\ & E \leq 2X+2M-N; \\ 2 \left(\frac{N-X-M-U-\frac{\delta}{2}}{N} \right), & N-U-\frac{\delta}{2} > X+M \geq \frac{N}{2}, \\ & E > 2X+2M-N; \\ \frac{N-2U-E}{N}, & X+M < \frac{N}{2}, \\ & U+E < X+M, \\ & N > 2U+E; \\ \frac{N-X-U-M}{N}, & X+M < \frac{N}{2}, \\ & U+E > X+M, \\ & N > X+M+U, \end{cases} \quad (12)$$

where $M = \max\{E, T\}$, and $\delta = N + E - 2X - 2M$.

Theorem 2: For a QSPIR system with N databases, where X of them can communicate to decode messages, B are Byzantine, E links are tapped by passive eavesdroppers that can listen to both uplink and downlink communications (static eavesdropper), U of the databases are unresponsive, and T of them are colluding to identify user-required message index, the following rate is achievable with symmetric privacy,

$$R_Q = \begin{cases} 2 \left(\frac{N-H-M-3B-U}{N} \right), & N-U > H+M+B \geq \frac{N}{2}, \\ & E \leq 2H+2M+2B-N; \\ 2 \left(\frac{N-H-M-3B-U-\frac{\delta}{2}}{N} \right), & N-U-\frac{\delta}{2} > H+M+B \geq \frac{N}{2}, \\ & E > 2H+2M+2B-N; \\ \frac{N-6B-2U-E}{N}, & H+M+B < \frac{N}{2}, \\ & 2B+U+E < H+M+B, \\ & N > 6B+2U+E; \\ \frac{N-H-3B-U-M}{N}, & H+M+B < \frac{N}{2}, \\ & 2B+U+E > H+M+B, \\ & N > H+3B+U+M, \end{cases} \quad (13)$$

where $M = \max\{E, T\}$, $H = \max\{X, B\}$, and $\delta = N + E - 2H - 2M - 2B$.

Theorem 3: For a QSPIR system with N databases, where X of them can communicate to decode messages, B are Byzantine, $|\mathcal{E}_1| \leq E$ of the uplinks and $|\mathcal{E}_2| \leq E$ of the downlinks are tapped by a passive eavesdropper, U of the databases are unresponsive, and T of them are colluding to

identify user-required message index, the following rate is achievable with symmetric privacy,

$$R_Q = \begin{cases} 2 \left(\frac{N-H-M-3B-U}{N} \right), & \begin{aligned} N-U > H+M+B \geq \frac{N}{2}, \\ E \leq 2H+2M+2B-N; \end{aligned} \\ 2 \left(\frac{N-H-M-3B-U-\frac{\delta}{2}}{N} \right), & \begin{aligned} N-U-\frac{\delta}{2} > H+M+B \geq \frac{N}{2}, \\ E > 2H+2M+2B-N; \end{aligned} \\ \frac{N-6B-2U-E}{N}, & \begin{aligned} H+M+B < \frac{N}{2}, \\ 2B+U+E < H+M+B, \\ N > 6B+2U+E; \end{aligned} \\ \frac{N-H-3B-U-M}{N}, & \begin{aligned} H+M+B < \frac{N}{2}, \\ 2B+U+E > H+M+B, \\ N > H+3B+U+2M, \end{aligned} \end{cases} \quad (14)$$

where $M = \max\{E+B, T\}$, $H = \max\{X, B\}$, and $\delta = N+E-2H-2M-2B$.

Remark 4: It is clear that there is a significant additional threat imposed on the system when we have both dynamic eavesdroppers and Byzantine servers in the system, while providing symmetric privacy. Further, allowing the eavesdropper to jump from any E links in the upload phase to another E links in the retrieval phase makes the setting more complex. This can be observed by comparing the rates in Theorem 1 and Theorem 3.

Remark 5: Note the difference in the effect of a dynamic eavesdropper in Theorem 3 compared to a static eavesdropper in Theorem 2 in terms of the achievable rate. The main difference lies in the fact that relaying the queries again by the Byzantine servers is not useful in the case of the static eavesdropper since they were able to fetch them during the uplink communication from the user to the servers.

Remark 6: In the achievable schemes sections, we present the achievable schemes for Theorem 1 and Theorem 3. The only modification required for the achievable scheme in Theorem 2 compared to that in Theorem 3 is to hide queries within a smaller vector subspace.

Remark 7: In Section VIII, we prove the properties of the main threat model given in Theorem 3, since other variations can be derived from it as they have fewer restrictions, i.e., $B=0$ for proving properties of Theorem 1, and $\mathcal{E}_1 = \mathcal{E}_2$ for Theorem 2.

Remark 8: In Theorem 2, the security parameter changes from X to $H = \max\{X, B\}$. Intuitively, the main reason behind this is to make sure that, if all B Byzantine servers send the m th symbol for the κ th message $W_{\kappa,m}$ to the user, the user will not be able to decode it.

Remark 9: In Theorem 3, the colluding parameter $M = \max\{E, T\}$ is modified to $M = \max\{E+B, T\}$. Logically, this is done to avoid the issue of Byzantine servers relaying the queries back and the potential of eavesdropping on those B links by the dynamic eavesdropper.

Remark 10: In both Theorem 2 and Theorem 3, a $3B$ factor appears instead of the usual $2B$ factor due to B Byzantine servers. The main reason for this is to hide the masking used for symmetric privacy from the collection of Byzantine servers. That is, if the masking variables are known, the

Byzantine servers can help the user to decode the interference terms and leak information about the other messages.

IV. PRELIMINARIES

The following notation is adopted in this paper to ease reading: For matrices, math typewriter font is used, e.g., $\mathbf{A}$. For column vectors, bold math font is used, e.g., $\mathbf{a}$.

A. Quantum Information Theory

We provide some important definitions for completeness. These can be found in introductory quantum information processing books, e.g., [42], [43].

Definition 1 (Quantum Density Matrices): For a quantum system A , that can be in the state $|\psi_j\rangle$ with probability p_j , the quantum density matrix ρ_A is defined as,

$$\rho_A = \sum_j p_j |\psi_j\rangle\langle\psi_j|, \quad (15)$$

with $p_j \geq 0$ and $\sum_j p_j = 1$.

Definition 2 (Von Neumann Entropy): For the density matrix ρ , the Von Neumann entropy is defined as,

$$S(\rho) = -\text{tr}(\rho \log \rho) = H(\Lambda), \quad (16)$$

where $\text{tr}(\cdot)$ is the trace operator, Λ are the eigenvalues of ρ , and $H(\cdot)$ is the Shannon entropy. For a quantum system A with density matrix ρ_A , we define $S(A) = S(\rho_A)$.

Definition 3 (Quantum Conditional Entropy): The conditional entropy of a quantum system A with respect to a system B is defined as,

$$S(A|B) = S(A, B) - S(B). \quad (17)$$

Definition 4 (Quantum Mutual Information): The quantum mutual information between two quantum systems A and B is defined as,

$$S(A; B) = S(A) + S(B) - S(A, B) \quad (18)$$

$$= S(A) - S(A|B). \quad (19)$$

Definition 5 (Quantum Operation): A quantum operation $\mathcal{E}$ is a linear, completely-positive, and trace-preserving map from the set of all density matrices to itself.

Definition 6 (Kraus Representation): Any quantum operation $\mathcal{E}$ acting on a quantum state ρ can be written in the form,

$$\mathcal{E}(\rho) = \sum_i M_i \rho M_i^\dagger, \quad (20)$$

where $\sum_i M_i^\dagger M_i = \mathbf{I}$.

B. N-Sum Box Abstraction

The encoding and decoding structure using the N -sum box abstraction introduced recently in [4] is defined as follows. In the encoding stage, the databases use generalized Pauli operators $\mathbf{X}(a) = \sum_{j=0}^{q-1} |j+a\rangle\langle j|$, and $\mathbf{Z}(a) = \sum_{j=0}^{q-1} \omega^{tr(aj)} |j\rangle\langle j|$, where $q = p^r$ with p being a prime number, $a \in \mathbb{F}_q$ and $\omega = \exp(2\pi i/p)$.

In the decoding stage, the user applies projection-valued measurement (PVM) defined on the quotient space of the stabilizer group $\mathcal{L}(\mathcal{V})$ defined by,

$$\mathcal{L}(\mathcal{V}) = \{c_v W(\mathbf{v}) : \mathbf{v} \in \mathcal{V}\}, \quad (21)$$

where $\mathcal{V}$ is a self-orthogonal subspace in $\mathbb{F}_q^{2N}$,

$$W(\mathbf{v}) = X(v_1)Z(v_{N+1}) \otimes \dots \otimes X(v_N)Z(v_{2N}), \quad (22)$$

and $c_v \in \mathbb{C}$ is chosen such that $\mathcal{L}(\mathcal{V})$ is an Abelian subgroup of HW_q^N with $c_v I_{q^N}$ being an element of the stabilizer group only with $c_v = 1$, where HW_q^N is the Heisenberg-Weyl group defined as follows,

$$HW_q^N = \{cW(s) : s \in \mathbb{F}_q^{2N}, c \in \mathbb{C} \setminus \{0\}\}. \quad (23)$$

Theorem 4 (Thm. 1 in [4]): Let G be a $2N \times N$ SSO matrix, i.e., $G^t J G = 0$, where $J = \begin{bmatrix} 0 & I_N \\ -I_N & 0 \end{bmatrix}$, and H be a $2N \times N$ matrix with full column rank. If $[G \ H]$ is a full-rank matrix, then $G' = [0_N \ I_N][G \ H]^{-1}$ is a feasible N -sum box transfer matrix.

Definition 7 (Dual QCSA Matrices [44]): The matrices H_N^u and H_N^v are defined as $H_N^u = \text{QCSA}_{N \times N}(\mathbf{f}_{[L]}, \alpha_{[N]}, \mathbf{u}_{[N]})$ and $H_N^v = \text{QCSA}_{N \times N}(\mathbf{f}_{[L]}, \alpha_{[N]}, \mathbf{v}_{[N]})$. Then, H_N^u and H_N^v are dual QCSA matrices if,

- 1) $u_1, \dots, u_N$ are non-zero,
- 2) $u_1, \dots, u_N$ are distinct,
- 3) for each v_j , $j \in [N]$,

$$v_j = \frac{1}{u_j} \left(\prod_{\substack{i=1 \\ i \neq j}}^N (\alpha_j - \alpha_i) \right)^{-1}, \quad (24)$$

where

$$\begin{aligned} \text{QCSA}_{N,L,N}(\mathbf{f}_{[L]}, \alpha_{[N]}, \beta_{[N]}) \\ = \text{diag}(\beta_{[N]}) \text{CSA}_{N,L,N}(\mathbf{f}_{[L]}, \alpha_{[N]}). \end{aligned} \quad (25)$$

Theorem 5 (Thm. 1 in [44]): For any dual QCSA matrices H_N^u and H_N^v , there exists a feasible N -sum box transfer matrix $G(\mathbf{u}, \mathbf{v})$ of size $N \times 2N$ given by,

$$G(\mathbf{u}, \mathbf{v}) = \begin{bmatrix} I_L & 0_{L \times \nu} & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{\mu-L} & 0 & 0 & 0 \\ 0 & 0 & 0 & I_L & 0_{L \times \mu} & 0 \\ 0 & 0 & 0 & 0 & 0 & I_{\nu-L} \end{bmatrix} \begin{bmatrix} H_N^u & 0 \\ 0 & H_N^v \end{bmatrix}^{-1}, \quad (26)$$

where $\nu = \lceil N/2 \rceil$ and $\mu = \lfloor N/2 \rfloor$.

V. ACHIEVABLE SCHEME FOR QXEUTSPIR

In this section and the following section, we follow the same encoding and decoding procedure for the N -sum box. In the naming of the subsections, we include the main restrictions for each regime, and trim the ones that basically state the requirement of having sufficiently many servers for the scheme to work.

A. *Regime 1: $X + M \geq \frac{N}{2}$ and $E \leq 2X + 2M - N$*

Let $L = N - X - M - U$ and the storage at the n th database, $S_n(i)$, $i \in [2]$, is

$$S_n(i) = \begin{bmatrix} \mathbf{W}_{\cdot,1}(i) + \sum_{j=1}^X (f_1 - \alpha_n)^j \mathbf{R}_{1j}(i) \\ \mathbf{W}_{\cdot,2}(i) + \sum_{j=1}^X (f_2 - \alpha_n)^j \mathbf{R}_{2j}(i) \\ \vdots \\ \mathbf{W}_{\cdot,L}(i) + \sum_{j=1}^X (f_L - \alpha_n)^j \mathbf{R}_{Lj}(i) \end{bmatrix}, \quad (27)$$

where $\mathbf{W}_{\cdot,j}(i) = [W_{1,j}(i), \dots, W_{K,j}(i)]^t$ is a vector representing the j th dit of all K messages, with $W_{k,j}$ being the j th bit of message k for the i th instance, $\mathbf{R}_{k,j}(i)$ are uniform independent random vectors with the same dimensions as $\mathbf{W}_{\cdot,j}(i)$, $j \in [1 : L]$, $\alpha_\ell, f_j \in \mathbb{F}_q$ are distinct where $\ell = 1, \dots, N$ and $j = 1, \dots, L$, and t denotes the transpose operator.

To retrieve the θ th message, the user sends the following queries to the n th server

$$\mathbf{Q}_n^{[\theta]} = \begin{bmatrix} \frac{1}{f_1 - \alpha_n} \left(\mathbf{e}_\theta + \sum_{j=1}^M (f_1 - \alpha_n)^j \mathbf{Z}_{1j} \right) \\ \vdots \\ \frac{1}{f_L - \alpha_n} \left(\mathbf{e}_\theta + \sum_{j=1}^M (f_L - \alpha_n)^j \mathbf{Z}_{Lj} \right) \end{bmatrix}, \quad (28)$$

where $\mathbf{e}_\theta$ is a vector of length K with 1 in the θ th index and zero otherwise, and $\mathbf{Z}_{ij}$ are uniform independent random vectors of length K each, chosen by the user.

Let $\mathcal{U} = \{n_1, \dots, n_U\}$ denote the set of unresponsive servers. The answer from the n th server, where $n \in [N] \setminus \mathcal{U}$ is given by

$$A_n^{[\theta]}(i) = S_n(i)^t \mathbf{Q}_n^{[\theta]} + \sum_{j=0}^{X+M-1} \alpha_n^j Z'_j(i), \quad (29)$$

where $Z'_j(i)$ are generated uniformly at random and shared among servers. Note that the answers from the $N - U$ servers, $i_1, \dots, i_{N-U}$, for one instance only of the classical scheme, can be written compactly as

$$\begin{bmatrix} A_{i_1}^{[\theta]} \\ \vdots \\ A_{i_{N-U}}^{[\theta]} \end{bmatrix} = \text{CSA}_{N-U,L,N-U} \mathbf{x}, \quad (30)$$

where

$$\mathbf{x} = \begin{bmatrix} \mathbf{W}_\theta \\ \mathbf{I}_{[X+M]} \end{bmatrix}, \quad (31)$$

and

$$\begin{aligned} \text{CSA}_{N-U,L,N-U} = \\ \begin{bmatrix} \frac{1}{f_1 - \alpha_{i_1}} & \dots & \frac{1}{f_L - \alpha_{i_1}} & 1 & \alpha_{i_1} & \dots & \alpha_{i_1}^{N-U-L-1} \\ \frac{1}{f_1 - \alpha_{i_2}} & \dots & \frac{1}{f_L - \alpha_{i_2}} & 1 & \alpha_{i_2} & \dots & \alpha_{i_2}^{N-U-L-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{1}{f_1 - \alpha_{i_{N-U}}} & \dots & \frac{1}{f_L - \alpha_{i_{N-U}}} & 1 & \alpha_{i_{N-U}} & \dots & \alpha_{i_{N-U}}^{N-U-L-1} \end{bmatrix}, \end{aligned} \quad (32)$$

where $\mathbf{I}_{[X+M]}$ is the interference vector which is discarded after decoding. In the classical case, the user does not receive any symbols from the unresponsive servers which can be identified easily in the classical regime. However, in the quantum setting the user has to apply initial measurements to know the unresponsive servers. Thus, in the quantum setting,

unresponsiveness can be thought of as an erasure error instead of non-received qudits as shown in [36]. The main reason for this is that the user can apply initial measurements to know which servers responded without affecting the remaining quantum state [45]. Based on the initial measurements, and the unresponsive servers, the actual measurements for decoding the message symbols can then be applied. As presented in [44], in the quantum setting two instances of the classical scheme can be transmitted in one instance using the quantum scheme when the servers encode the symbols on the entangled shared state whenever possible. Thus, the answers of the responsive servers can be written for every $n \in [N] \setminus \mathcal{U}$ as

$$X(u_n A_n(1))Z(v_n A_n(2))\rho_n (X(u_n A_n(1))Z(v_n A_n(2)))^\dagger, \quad (33)$$

where ρ_n is the density matrix for the quantum system $\mathcal{A}_n$.

As for the unresponsive servers in the quantum setting, the answers for $n \in \mathcal{U}$ can be written as

$$\begin{aligned} & X(u_n \Delta'_n(1))Z(v_n \Delta'_n(2))X(u_n A_n(1))Z(v_n A_n(2))\rho_n \\ & \times (X(u_n \Delta'_n(1))Z(v_n \Delta'_n(2))X(u_n A_n(1))Z(v_n A_n(2)))^\dagger \\ & = X(u_n (A_n(1) + \Delta'_n(1)))Z(v_n (A_n(2) + \Delta'_n(2)))\rho_n \\ & \times (X(u_n (A_n(1) + \Delta'_n(1)))Z(v_n (A_n(2) + \Delta'_n(2))))^\dagger, \end{aligned} \quad (34)$$

where $\mathcal{A}_n$ is the subsystem available at the n th database, Δ'_n is generated uniformly at random, and $A_n(i)$ is as in (29).

Now, for one instance of the quantum scheme, the collective answers can be written as follows

$$A^{[\theta]} = \begin{bmatrix} \Lambda_{11} & 0 & \Gamma_1 & 0 & \Lambda_{12} & 0 & \Omega & 0 \\ 0 & \Lambda_{21} & 0 & \Gamma_2 & 0 & \Lambda_{22} & 0 & \Omega \end{bmatrix} \begin{bmatrix} I_1(1) \\ I_1(2) \\ \mathbf{W}_\theta(1) \\ \mathbf{W}_\theta(2) \\ I_2(1) \\ I_2(2) \\ \Delta'_U(1) \\ \Delta'_U(2) \end{bmatrix}, \quad (35)$$

where $\Lambda_{11} = \text{GRS}_{N,X+M}^{\alpha_{[N]},u_{[N]}}(:, 1 : \lceil \frac{N}{2} \rceil)$, $\Lambda_{12} = \text{GRS}_{N,X+M}^{\alpha_{[N]},u_{[N]}}(:, \lceil \frac{N}{2} \rceil + 1 : X + M)$, $\Lambda_{21} = \text{GRS}_{N,X+M}^{\alpha_{[N]},v_{[N]}}(:, 1 : \lfloor \frac{N}{2} \rfloor)$, $\Lambda_{22} = \text{GRS}_{N,X+M}^{\alpha_{[N]},v_{[N]}}(:, \lfloor \frac{N}{2} \rfloor + 1 : X + M)$, $\Gamma_1 = \text{GC}_{N,L}^{\alpha_{[N]},u_{[N]}}(f_{[L]})$, $\Gamma_2 = \text{GC}_{N,L}^{\alpha_{[N]},v_{[N]}}(f_{[L]})$, and $\Omega = [\mathbf{e}_{n_1}, \mathbf{e}_{n_2}, \dots, \mathbf{e}_{n_U}]$, where $\mathbf{e}_{n_i}$ is the unit vector of length N and 0 in all locations except the n_i th index, and $\mathcal{U} = \{n_1, \dots, n_U\}$.

In addition,

$$\text{GRS}_{N,X+M}^{\alpha_{[N]},u_{[N]}} = \begin{bmatrix} u_1 & u_1 \alpha_1 & \dots & u_1 \alpha_1^{X+M-1} \\ u_2 & u_2 \alpha_2 & \dots & u_2 \alpha_2^{X+M-1} \\ \vdots & \vdots & \ddots & \vdots \\ u_N & u_N \alpha_N & \dots & u_N \alpha_N^{X+M-1} \end{bmatrix}, \quad (36)$$

and

$$\text{GC}_{N,L}^{\alpha_{[N]},u_{[N]}}(f_{[L]}) = \begin{bmatrix} \frac{u_1}{f_1 - \alpha_1} & \frac{u_1}{f_2 - \alpha_1} & \dots & \frac{u_1}{f_L - \alpha_1} \\ \frac{u_2}{f_1 - \alpha_2} & \frac{u_2}{f_2 - \alpha_2} & \dots & \frac{u_2}{f_L - \alpha_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{u_N}{f_1 - \alpha_N} & \frac{u_N}{f_2 - \alpha_N} & \dots & \frac{u_N}{f_L - \alpha_N} \end{bmatrix}. \quad (37)$$

Since there are $E \leq 2X + 2M - N$ eavesdropped links, the servers can use a channel transmission matrix as follows

$$G'(\mathbf{u}, \mathbf{v}) = \underbrace{\begin{bmatrix} \mathbf{I}_N & 0 & 0 \\ 0 & \tilde{\mathbf{V}}_1 & 0 \\ 0 & 0 & \mathbf{I}_{2U} \end{bmatrix}}_{\mathbf{M}_T^{-1}} \mathbf{M}^{-1}, \quad (38)$$

where $\tilde{\mathbf{V}}_1 = \mathbf{V}_{2L+2X+2M-N}(b_{[2L+2X+2M-N]})$, and in general we denote

$$\mathbf{V}_N(\mathbf{b}_{[N]}) = \begin{bmatrix} 1 & b_1 & \dots & b_1^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & b_N & \dots & b_N^{N-1} \end{bmatrix}, \quad (39)$$

and

$$\mathbf{M} = \begin{bmatrix} \Lambda_{11} & 0 & \Gamma_1 & 0 & \Lambda_{12} & 0 & \Omega & 0 \\ 0 & \Lambda_{21} & 0 & \Gamma_2 & 0 & \Lambda_{22} & 0 & \Omega \end{bmatrix}. \quad (40)$$

Remark 11: The proof for the feasibility of the chosen transmission matrix $G'(\mathbf{u}, \mathbf{v})$ for this regime and the remaining regimes is given in Section VIII.

Now, the received symbols by the user after the final measurements can be written as

$$\mathbf{y} = \begin{bmatrix} \mathbf{I}_N & 0 & 0 \\ 0 & \tilde{\mathbf{V}}_1 & 0 \\ 0 & 0 & \mathbf{I}_{2U} \end{bmatrix} \begin{bmatrix} I_1(1) \\ I_1(2) \\ \mathbf{W}_\theta(1) \\ \mathbf{W}_\theta(2) \\ I_2(1) \\ I_2(2) \\ \Delta'_U(1) \\ \Delta'_U(2) \end{bmatrix} \quad (41)$$

$$= \begin{bmatrix} \tilde{\mathbf{V}}_1 & \begin{bmatrix} \mathbf{W}_\theta(1) \\ \mathbf{W}_\theta(2) \\ I_2(1) \\ I_2(2) \\ \Delta'_U(1) \\ \Delta'_U(2) \end{bmatrix} \end{bmatrix}. \quad (42)$$

Since the first submatrix in (42) is invertible, the user can decode the message symbols.

B. Regime 2: $X + M \geq \frac{N}{2}$ and $E > 2X + 2M - N$

In this case, we develop a scheme that can use the interference symbols along with the noise variables introduced in the storage to prevent the eavesdropper from decoding the messages. As the number of eavesdropped links is greater than the number of interference symbols received by the user, we need to introduce new noise variables to compensate for the difference. Let $\delta = E - (2X + 2M - N)$, $L_1 = N - X - M - U - \delta$, and $L_2 = L_1 + \delta$. Define $\mathbf{r}_{\cdot,1}, \dots, \mathbf{r}_{\cdot,\delta}$ to be uniform random vectors in $\mathbb{F}_q$. Then, the storage is given by

$$\mathbf{S}_n(1) = \begin{bmatrix} \mathbf{r}_{\cdot,1} + \sum_{j=1}^X (f_1 - \alpha_n)^j \mathbf{R}_{1,j}(1) \\ \vdots \\ \mathbf{r}_{\cdot,\delta} + \sum_{j=1}^X (f_\delta - \alpha_n)^j \mathbf{R}_{\delta,j}(1) \\ \mathbf{W}_{\cdot,1} + \sum_{j=1}^X (f_{\delta+1} - \alpha_n)^j \mathbf{R}_{\delta+1,j}(1) \\ \mathbf{W}_{\cdot,2} + \sum_{j=1}^X (f_{\delta+2} - \alpha_n)^j \mathbf{R}_{\delta+2,j}(1) \\ \vdots \\ \mathbf{W}_{\cdot,L_1} + \sum_{j=1}^X (f_{L_2} - \alpha_n)^j \mathbf{R}_{L_1+\delta,j}(1) \end{bmatrix}, \quad (43)$$

and

$$\mathbf{S}_n(2) = \begin{bmatrix} \mathbf{W}_{:,L_1+1} + \sum_{j=1}^X (f_1 - \alpha_n)^j \mathbf{R}_{1j}(2) \\ \mathbf{W}_{:,L_1+2} + \sum_{j=1}^X (f_2 - \alpha_n)^j \mathbf{R}_{2j}(2) \\ \vdots \\ \mathbf{W}_{:,L_1+L_2} + \sum_{j=1}^X (f_{L_2} - \alpha_n)^j \mathbf{R}_{L_2,j}(2) \end{bmatrix}. \quad (44)$$

Using the same procedure as in the previous section, the user sends the following query to the n th server

$$\mathbf{Q}_n^{[\theta]} = \begin{bmatrix} \frac{1}{f_1 - \alpha_n} \left(\mathbf{e}_\theta + \sum_{j=1}^M (f_1 - \alpha_n)^j \mathbf{Z}_{1j} \right) \\ \vdots \\ \frac{1}{f_{L_2} - \alpha_n} \left(\mathbf{e}_\theta + \sum_{j=1}^M (f_{L_2} - \alpha_n)^j \mathbf{Z}_{L_2,j} \right) \end{bmatrix}. \quad (45)$$

The transmission matrix is now slightly modified as follows

$$\mathbf{G}'(\mathbf{u}, \mathbf{v}) = \begin{bmatrix} \mathbf{I}_N & 0 & 0 \\ 0 & \tilde{\mathbf{V}}_2 & 0 \\ 0 & 0 & \mathbf{I}_{2U} \end{bmatrix} \underbrace{\mathbf{M}^{-1}}_{\mathbf{M}_2^{-1}}, \quad (46)$$

where $\tilde{\mathbf{V}}_2 = \mathbf{V}_{2L_2+2X+2M-N}(\mathbf{b}_{[2L_2+2X+2M-N]})$, and $\mathbf{M}$ is the same as (40).

The received symbols at the user side can be derived as follows

$$\mathbf{y} = \begin{bmatrix} \mathbf{I}_N & 0 & 0 \\ 0 & \tilde{\mathbf{V}}_2 & 0 \\ 0 & 0 & \mathbf{I}_{2U} \end{bmatrix} \begin{bmatrix} \mathbf{I}_1(1) \\ \mathbf{I}_1(2) \\ \mathbf{r}_{[\delta]}^{[\theta]} \\ \mathbf{W}_\theta(1) \\ \mathbf{W}_\theta(2) \\ \mathbf{I}_2(1) \\ \mathbf{I}_2(2) \\ \Delta'_U(1) \\ \Delta'_U(2) \end{bmatrix} \quad (47)$$

$$= \begin{bmatrix} \tilde{\mathbf{V}}_2 \begin{bmatrix} \mathbf{r}_{[\delta]}^{[\theta]} \\ \mathbf{W}_\theta(1) \\ \mathbf{W}_\theta(2) \\ \mathbf{I}_2(1) \\ \mathbf{I}_2(2) \end{bmatrix} \\ \Delta'_U(1) \\ \Delta'_U(2) \end{bmatrix}. \quad (48)$$

Thus, since the first instance has L_1 useful symbols of the required message symbols and the second instance has L_2 useful symbols, the rate is given by $R_Q = \frac{L_1+L_2}{N} = 2 \left(\frac{N-X-M-U-\frac{\delta}{2}}{N} \right)$.

C. Regime 3: $U + E < X + M < \frac{N}{2}$

In this regime, the user chooses $T_1 \geq T$ and $T_2 \geq T$ such that $T_1 + T_2 + 2X = N$ and $T_1 \leq T_2 \leq T_1 + 1$. Let $L_1 = \lfloor \frac{N}{2} \rfloor - U - E$

and $L_2 = \lfloor \frac{N}{2} \rfloor - U$. Thus, the storage is given by

$$\mathbf{S}_n(1) = \begin{bmatrix} \mathbf{r}_{:,1} + \sum_{j=1}^X (f_1 - \alpha_n)^j \mathbf{R}_{1j}(1) \\ \vdots \\ \mathbf{r}_{:,E} + \sum_{j=1}^X (f_E - \alpha_n)^j \mathbf{R}_{Ej}(1) \\ \mathbf{W}_{:,1} + \sum_{j=1}^X (f_{E+1} - \alpha_n)^j \mathbf{R}_{E+1,j}(1) \\ \mathbf{W}_{:,2} + \sum_{j=1}^X (f_{E+2} - \alpha_n)^j \mathbf{R}_{E+2,j}(1) \\ \vdots \\ \mathbf{W}_{:,L_1} + \sum_{j=1}^X (f_{L_2} - \alpha_n)^j \mathbf{R}_{L_1+E,j}(1) \end{bmatrix}, \quad (49)$$

and

$$\mathbf{S}_n(2) = \begin{bmatrix} \mathbf{W}_{:,L_1+1} + \sum_{j=1}^X (f_1 - \alpha_n)^j \mathbf{R}_{1j}(2) \\ \mathbf{W}_{:,L_1+2} + \sum_{j=1}^X (f_2 - \alpha_n)^j \mathbf{R}_{2j}(2) \\ \vdots \\ \mathbf{W}_{:,L_1+L_2} + \sum_{j=1}^X (f_{L_2} - \alpha_n)^j \mathbf{R}_{L_2,j}(2) \end{bmatrix}. \quad (50)$$

The queries are given by

$$\mathbf{Q}_n^{[\theta]}(1) = \begin{bmatrix} \frac{1}{f_1 - \alpha_n} \left(\mathbf{e}_\theta + \sum_{j=1}^{T_1} (f_1 - \alpha_n)^j \mathbf{Z}_{1j}(1) \right) \\ \vdots \\ \frac{1}{f_{L_2} - \alpha_n} \left(\mathbf{e}_\theta + \sum_{j=1}^{T_1} (f_{L_2} - \alpha_n)^j \mathbf{Z}_{L_2,j}(1) \right) \end{bmatrix}, \quad (51)$$

and

$$\mathbf{Q}_n^{[\theta]}(2) = \begin{bmatrix} \frac{1}{f_1 - \alpha_n} \left(\mathbf{e}_\theta + \sum_{j=1}^{T_2} (f_1 - \alpha_n)^j \mathbf{Z}_{1j}(2) \right) \\ \vdots \\ \frac{1}{f_{L_2} - \alpha_n} \left(\mathbf{e}_\theta + \sum_{j=1}^{T_1} (f_{L_2} - \alpha_n)^j \mathbf{Z}_{L_2,j}(2) \right) \end{bmatrix}. \quad (52)$$

Each answer instance from the n th server is given by

$$\mathbf{A}_n^{[\theta]}(i) = \mathbf{S}_n(i)^t \mathbf{Q}_n^{[\theta]}(i) + \sum_{j=1}^{X+T_i} \alpha_n^{i-1} \mathbf{Z}_j'(i), \quad i \in [2]. \quad (53)$$

The transmission matrix for this regime is the same as the previous section with $\delta = E$, and $2X + 2M = N$, i.e.,

$$\mathbf{G}'(\mathbf{u}, \mathbf{v}) = \begin{bmatrix} \mathbf{I}_N & 0 & 0 \\ 0 & \tilde{\mathbf{V}}_3 & 0 \\ 0 & 0 & \mathbf{I}_{2U} \end{bmatrix} \underbrace{\mathbf{M}^{-1}}_{\mathbf{M}_2^{-1}}, \quad (54)$$

where $\tilde{\mathbf{V}}_3 = \mathbf{V}_{L_1+L_2+E}(\mathbf{b}_{[L_1+L_2+E]})$. Thus, the received symbols at the user side can be derived as follows

$$\mathbf{y} = \begin{bmatrix} \mathbf{I}_N & 0 & 0 \\ 0 & \tilde{\mathbf{V}}_3 & 0 \\ 0 & 0 & \mathbf{I}_{2U} \end{bmatrix} \begin{bmatrix} \mathbf{I}_1(1) \\ \mathbf{I}_1(2) \\ \mathbf{r}_{[E]}^{[\theta]} \\ \mathbf{W}_\theta(1) \\ \mathbf{W}_\theta(2) \\ \Delta'_U(1) \\ \Delta'_U(2) \end{bmatrix} \quad (55)$$

$$= \begin{bmatrix} \tilde{\mathbf{V}}_3 \begin{bmatrix} \mathbf{r}_{[E]}^{[\theta]} \\ \mathbf{W}_\theta(1) \\ \mathbf{W}_\theta(2) \end{bmatrix} \\ \Delta'_U(1) \\ \Delta'_U(2) \end{bmatrix}. \quad (56)$$

Then, by applying the same calculation steps, the rate is $R_Q = \frac{L_1+L_2}{N} = \frac{N-2U-E}{N}$.

Remark 12: If $T_1 = T_2$, then it is sufficient to send one query instance only to the servers.

Remark 13: Note that $T_1 \geq E$ and $T_2 \geq E$, since the following inequalities hold

$$X + M < \frac{N}{2} \Leftrightarrow X + M \leq \lfloor \frac{N}{2} \rfloor \quad (57)$$

$$\Leftrightarrow X + T \leq \lfloor \frac{N}{2} \rfloor \quad (58)$$

$$\Leftrightarrow 2X + 2E \leq 2\lfloor \frac{N}{2} \rfloor \quad (59)$$

$$\Rightarrow 2E \leq T_1 + T_2 \quad (60)$$

$$\Leftrightarrow 2E \leq 2T_1 + 1 \quad (61)$$

$$\Rightarrow E \leq T_1 \quad (62)$$

$$\Rightarrow E \leq T_2, \quad (63)$$

where for (58), we used $X + E \leq \lfloor \frac{N}{2} \rfloor$ and (60) is due to the fact $2X + T_1 + T_2 = N$.

D. Regime 4: $X + M < \frac{N}{2}$ and $U + E > X + M$

In this setting, the usage of the classical scheme is preferable compared to the quantum scheme since there are not enough dimensions for transmitting message symbols using the quantum scheme or the rate of the classical scheme is better than the quantum scheme. For the sake of completeness, we present the classical scheme here. The sub-packetization length is $L = N - X - M - U$ and the storage is as in (27) and queries by the user are as in (28). In addition, the answers are computed as

$$A_n^{[\theta]} = \mathbf{S}_n^t \mathbf{Q}_n^{[\theta]} + \sum_{i=1}^{X+M} \alpha_n^{i-1} Z'_i, \quad n \in [N] \setminus \mathcal{U}. \quad (64)$$

Thus, the answers can be written collectively as

$$\begin{bmatrix} A_{i_1}^{[\theta]} \\ \vdots \\ A_{i_{N-U}}^{[\theta]} \end{bmatrix} = \mathbf{CSA}_{N-U, L, N-U} \begin{bmatrix} W_{\theta,1} \\ \vdots \\ W_{\theta,L} \\ I_1 + Z'_1 \\ \vdots \\ I_{X+M} + Z'_{X+M} \end{bmatrix}, \quad (65)$$

where

$$\mathbf{CSA}_{N-U, L, N-U} = \begin{bmatrix} \frac{1}{f_1 - \alpha_{i_1}} & \cdots & \frac{1}{f_L - \alpha_{i_1}} & 1 & \alpha_{i_1} & \cdots & \alpha_{i_1}^{X+M-1} \\ \frac{1}{f_1 - \alpha_{i_2}} & \cdots & \frac{1}{f_L - \alpha_{i_2}} & 1 & \alpha_{i_2} & \cdots & \alpha_{i_2}^{X+M-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{1}{f_1 - \alpha_{i_{N-U}}} & \cdots & \frac{1}{f_L - \alpha_{i_{N-U}}} & 1 & \alpha_{i_{N-U}} & \cdots & \alpha_{i_{N-U}}^{X+M-1} \end{bmatrix} \quad (66)$$

VI. ACHIEVABLE SCHEME FOR QXBEUTSPIR

In this section, we only analyze the actions of the Byzantine servers on the N -sum box as a black box, i.e., they use the same quantum operations the honest servers use but can manipulate the transmitted dits. To extend this to all

arbitrary quantum operations, we use the fact that any quantum operation can be written as a linear combination of the Pauli operators. Thus, the achievable rates in this scheme will not change even if the Byzantine servers apply any quantum operations on their transmitted dits. The concrete analysis for this observation can be found in [46].

A. Regime 1: $H + M + B \geq \frac{N}{2}$ and $E \leq 2H + 2M + 2B - N$

Let $L = N - H - M - 3B - U$ and the storage at the n th database, $\mathbf{S}_n(i)$, $i \in [2]$,

$$\mathbf{S}_n(i) = \begin{bmatrix} \mathbf{W}_{\cdot,1}(i) + \sum_{j=1}^H (f_1 - \alpha_n)^j \mathbf{R}_{1j}(i) \\ \mathbf{W}_{\cdot,2}(i) + \sum_{j=1}^H (f_2 - \alpha_n)^j \mathbf{R}_{2j}(i) \\ \vdots \\ \mathbf{W}_{\cdot,L}(i) + \sum_{j=1}^H (f_L - \alpha_n)^j \mathbf{R}_{Lj}(i) \end{bmatrix}. \quad (67)$$

In addition, the noise used to provide symmetric privacy and protection against collaboration between the eavesdropper and the Byzantine servers is distributed to all servers as follows

$$\hat{Z}_n(i) = Z'_1(i) + \alpha_n Z'_2(i) + \cdots + \alpha_n^{M+H-1} Z'_{H+M}(i) + \alpha_n^{M+H} R'_1(i) + \cdots + \alpha_n^{M+H+B-1} R'_B(i), \quad (68)$$

where $\{Z'_j, R'_j\}$ are uniform independent random variables. The queries are given by

$$\mathbf{Q}_n^{[\theta]} = \begin{bmatrix} \frac{1}{f_1 - \alpha_n} (\mathbf{e}_\theta + \sum_{j=1}^M (f_1 - \alpha_n)^j \mathbf{Z}_{1j}) \\ \vdots \\ \frac{1}{f_L - \alpha_n} (\mathbf{e}_\theta + \sum_{j=1}^M (f_L - \alpha_n)^j \mathbf{Z}_{Lj}) \end{bmatrix}, \quad (69)$$

where $\mathbf{e}_\theta$ is a vector of length K with 1 in the θ th index and zero otherwise, and $\mathbf{Z}_{ij}$ are uniform independent random vectors of length K each, chosen by the user.

Let $\mathcal{B} = \{n_{i_1}, \dots, n_{i_B}\}$ denote the indices of the Byzantine servers, and as in the previous section $\mathcal{U}$ denotes the indices of the unresponsive servers. Thus, the answer from the n th server, $n \in [N] \setminus (\mathcal{B} \cup \mathcal{U})$, is given by

$$A_n^{[\theta]}(i) = \mathbf{S}_n(i)^t \mathbf{Q}_n^{[\theta]} + \hat{Z}_n(i), \quad (70)$$

and for $n \in \mathcal{B}$ is given by

$$A_n^{[\theta]}(i) = \eta_n(i) = \mathbf{S}_n(i)^t \mathbf{Q}_n^{[\theta]} + \hat{Z}_n(i) + \Delta_n(i). \quad (71)$$

The answers from the $N - U$ servers, $i_1, \dots, i_{N-U}$, for one instance, can be written collectively as shown in [37], i.e.,

$$\begin{bmatrix} A_{i_1}^{[\theta]} \\ \vdots \\ A_{i_{N-U}}^{[\theta]} \end{bmatrix} = \mathbf{CSA}_{N-U, L, N-U} \begin{bmatrix} \mathbf{W}_\theta \\ \mathbf{I}_{[H+M]} + \mathbf{Z}'_{[H+M]} \\ \mathbf{R}'_{[B]} \\ \mathbf{0}_{2B} \end{bmatrix} + \Delta_B \quad (72)$$

$$= \mathbf{CSA}_{N-U, L, N-U} \mathbf{x}, \quad (73)$$

where

$$\mathbf{x} = \begin{bmatrix} \mathbf{W}_\theta \\ \mathbf{I}_{[H+M]} + \mathbf{Z}'_{[H+M]} \\ \mathbf{R}'_{[B]} \\ \mathbf{0}_{2B} \end{bmatrix} + \mathbf{CSA}_{N-U, L, N-U}^{-1} \Delta_B, \quad (74)$$

and Δ_B is a vector of length $N - U$ with only B non-zero elements.

Based on the analysis in the previous section about the unresponsive servers, the two answer instances can be written as follows

$$\mathbf{A}^{[\theta]} = \underbrace{\begin{bmatrix} \Lambda'_{11} & 0 & \Gamma'_1 & 0 & \Lambda'_{12} & 0 & \Omega & 0 \\ 0 & \Lambda'_{21} & 0 & \Gamma'_2 & 0 & \Lambda'_{22} & 0 & \Omega \end{bmatrix}}_{\mathbf{M}'} \begin{bmatrix} \tilde{\mathbf{I}}_1(1) \\ \tilde{\mathbf{I}}_1(2) \\ \tilde{\mathbf{W}}_\theta(1) \\ \tilde{\mathbf{W}}_\theta(2) \\ \tilde{\mathbf{I}}_2(1) \\ \tilde{\mathbf{I}}_2(2) \\ \tilde{\Delta}_B(1) \\ \tilde{\Delta}_B(2) \\ \Delta'_U(1) \\ \Delta'_U(2) \end{bmatrix}, \quad (75)$$

where $\Lambda'_{11} = \text{GRS}_{N,H+M}^{\alpha_{[N]}, \mathbf{u}_{[N]}}(:, 1 : \lceil \frac{N}{2} \rceil)$, $\Lambda'_{12} = \text{GRS}_{N,H+M}^{\alpha_{[N]}, \mathbf{u}_{[N]}}(:, \lceil \frac{N}{2} \rceil + 1 : H + M + 3B)$, $\Lambda'_{21} = \text{GRS}_{N,H+M}^{\alpha_{[N]}, \mathbf{v}_{[N]}}(:, 1 : \lfloor \frac{N}{2} \rfloor)$, $\Lambda'_{22} = \text{GRS}_{N,H+M}^{\alpha_{[N]}, \mathbf{v}_{[N]}}(:, \lfloor \frac{N}{2} \rfloor + 1 : H + M + 3B)$, $\Gamma'_1 = \text{GC}_{N,L}^{\alpha_{[N]}, \mathbf{u}_{[N]}} \mathbf{f}_{[L]}$, $\Gamma'_2 = \text{GC}_{N,L}^{\alpha_{[N]}, \mathbf{v}_{[N]}} \mathbf{f}_{[L]}$, $\Omega = [\mathbf{e}_{n_1}, \mathbf{e}_{n_2}, \dots, \mathbf{e}_{n_U}]$, and

$$\begin{bmatrix} \tilde{\mathbf{W}}_\theta(i) \\ \tilde{\mathbf{I}}_1(i) \\ \tilde{\mathbf{I}}_2(i) \\ \tilde{\Delta}_B(i) \end{bmatrix} = \begin{bmatrix} \mathbf{W}_\theta(i) \\ \hat{\mathbf{I}}_{[H+M+B]}(i) \\ \mathbf{0}_{2B} \end{bmatrix} + \text{CSA}_{N-U,L,N-U}^{-1} \Delta_B(i), \quad (76)$$

where

$$\hat{\mathbf{I}}_{[H+M+B]}(i) = \begin{bmatrix} \mathbf{I}_{[H+M]} + \mathbf{Z}'_{[H+M]} \\ \mathbf{R}'_{[B]} \end{bmatrix}. \quad (77)$$

Since there are $E \leq 2H + 2M + 2B - N$ eavesdropped links, the servers can use a channel transmission matrix as follows

$$\mathbf{G}'(\mathbf{u}, \mathbf{v}) = \begin{bmatrix} \mathbf{I}_N & 0 & 0 \\ 0 & \tilde{\mathbf{V}}_3 & 0 \\ 0 & 0 & \mathbf{I}_{4B+2U} \end{bmatrix} \mathbf{M}'^{-1}, \quad (78)$$

where $\tilde{\mathbf{V}}_3 = \mathbf{V}_{2L+2H+2M+2B-N}(\mathbf{b}_{[2L+2H+2M+2B-N]})$, and $\mathbf{M}'$ is as shown in (75). Now, the received symbols can be written as

$$\mathbf{y} = \begin{bmatrix} 0_N & \mathbf{I}_N \end{bmatrix} \begin{bmatrix} \mathbf{I}_N & 0 & 0 \\ 0 & \tilde{\mathbf{V}}_3 & 0 \\ 0 & 0 & \mathbf{I}_{4B+2U} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{I}}_1(1) \\ \tilde{\mathbf{I}}_1(2) \\ \tilde{\mathbf{W}}_\theta(1) \\ \tilde{\mathbf{W}}_\theta(2) \\ \tilde{\mathbf{I}}_2(1) \\ \tilde{\mathbf{I}}_2(2) \\ \tilde{\Delta}_B(1) \\ \tilde{\Delta}_B(2) \\ \Delta'_U(1) \\ \Delta'_U(2) \end{bmatrix} \quad (79)$$

$$= \begin{bmatrix} \tilde{\mathbf{V}}_3 \begin{bmatrix} \tilde{\mathbf{W}}_\theta(1) \\ \tilde{\mathbf{W}}_\theta(2) \\ \tilde{\mathbf{I}}_2(1) \\ \tilde{\mathbf{I}}_2(2) \end{bmatrix} \\ \tilde{\Delta}_B(1) \\ \tilde{\Delta}_B(2) \\ \Delta'_U(1) \\ \Delta'_U(2) \end{bmatrix}. \quad (80)$$

In Section VII, we show the decoding scheme that is applied on (80) to decode the message symbols. Note that since $\mathbf{V}_{2L+2H+2M+2B-N}(\mathbf{b}_{[2L+2H+2M+2B-N]})$ is invertible, we have

$$\mathbf{y}_1 = \begin{bmatrix} \tilde{\mathbf{V}}_3^{-1} & 0 \\ 0 & \mathbf{I}_{4B+2U} \end{bmatrix} \mathbf{y} \quad (81)$$

$$= \begin{bmatrix} \tilde{\mathbf{W}}_\theta(1) \\ \tilde{\mathbf{W}}_\theta(2) \\ \tilde{\mathbf{I}}_2(1) \\ \tilde{\mathbf{I}}_2(2) \\ \tilde{\Delta}_B(1) \\ \tilde{\Delta}_B(2) \\ \Delta'_U(1) \\ \Delta'_U(2) \end{bmatrix}. \quad (82)$$

B. Regime 2: $H + M + B \geq \frac{N}{2}$ and $E > 2H + 2M + 2B - N$

In this case, we develop a scheme that can use the interference symbols along with the noise variables introduced in the storage to prevent the eavesdropper from decoding the messages. As the number of eavesdropped links is greater than the number of received interference symbols, we need to introduce new noise variables to compensate for the difference as mentioned previously. Let $\delta = E - (2H + 2M + 2B - N)$, $L_1 = N - H - M - 3B - U - \delta$, and $L_2 = L_1 + \delta$. Define $\mathbf{r}_{\cdot,1}, \dots, \mathbf{r}_{\cdot,\delta}$ as uniform random vectors. Then, the storage is given by

$$\mathbf{S}_n(1) = \begin{bmatrix} \mathbf{r}_{\cdot,1} + \sum_{j=1}^H (f_1 - \alpha_n)^j \mathbf{R}_{1,j}(1) \\ \vdots \\ \mathbf{r}_{\cdot,\delta} + \sum_{j=1}^H (f_\delta - \alpha_n)^j \mathbf{R}_{\delta,j}(1) \\ \mathbf{W}_{\cdot,1} + \sum_{j=1}^H (f_{\delta+1} - \alpha_n)^j \mathbf{R}_{\delta+1,j}(1) \\ \mathbf{W}_{\cdot,2} + \sum_{j=1}^H (f_{\delta+2} - \alpha_n)^j \mathbf{R}_{\delta+2,j}(1) \\ \vdots \\ \mathbf{W}_{\cdot,L_1} + \sum_{j=1}^H (f_{L_2} - \alpha_n)^j \mathbf{R}_{L_1+\delta,j}(1) \end{bmatrix}, \quad (83)$$

and

$$\mathbf{S}_n(2) = \begin{bmatrix} \mathbf{W}_{\cdot,L_1+1} + \sum_{j=1}^H (f_1 - \alpha_n)^j \mathbf{R}_{1,j}(2) \\ \mathbf{W}_{\cdot,L_1+2} + \sum_{j=1}^H (f_2 - \alpha_n)^j \mathbf{R}_{2,j}(2) \\ \vdots \\ \mathbf{W}_{\cdot,L_1+L_2} + \sum_{j=1}^H (f_{L_2} - \alpha_n)^j \mathbf{R}_{L_2,j}(2) \end{bmatrix}. \quad (84)$$

Using the same procedure as in the previous section, the rate is given by $R_Q = \frac{L_1+L_2}{N} = 2 \left(\frac{N-H-M-3B-U-\frac{\delta}{2}}{N} \right)$.

C. Regime 3: $2B + U + E < H + M + B < \frac{N}{2}$

In this regime, the user chooses $T_1 \geq T$ and $T_2 \geq T$ such that $T_1 + T_2 + 2H + 2B = N$ and $T_1 \leq T_2 \leq T_1 + 1$. Let $L_1 = \lceil \frac{N}{2} \rceil - 3B - U - E$ and $L_2 = \lfloor \frac{N}{2} \rfloor - 3B - U$. Thus, the storage is given by

$$\mathbf{S}_n(1) = \begin{bmatrix} \mathbf{r}_{\cdot,1} + \sum_{j=1}^H (f_1 - \alpha_n)^j \mathbf{R}_{1,j}(1) \\ \vdots \\ \mathbf{r}_{\cdot,E} + \sum_{j=1}^H (f_E - \alpha_n)^j \mathbf{R}_{E,j}(1) \\ \mathbf{W}_{\cdot,1} + \sum_{j=1}^H (f_{E+1} - \alpha_n)^j \mathbf{R}_{E+1,j}(1) \\ \mathbf{W}_{\cdot,2} + \sum_{j=1}^H (f_{E+2} - \alpha_n)^j \mathbf{R}_{E+2,j}(1) \\ \vdots \\ \mathbf{W}_{\cdot,L_1} + \sum_{j=1}^H (f_{L_2} - \alpha_n)^j \mathbf{R}_{L_1+E,j}(1) \end{bmatrix}, \quad (85)$$

and

$$S_n(2) = \begin{bmatrix} \mathbf{W}_{\cdot, L_1+1} + \sum_{j=1}^H (f_1 - \alpha_n)^j \mathbf{R}_{1j}(2) \\ \mathbf{W}_{\cdot, L_1+2} + \sum_{j=1}^H (f_2 - \alpha_n)^j \mathbf{R}_{2j}(2) \\ \vdots \\ \mathbf{W}_{\cdot, L_1+L_2} + \sum_{j=1}^H (f_{L_2} - \alpha_n)^j \mathbf{R}_{L_2,j}(2) \end{bmatrix}. \quad (86)$$

In addition, we have the same interference symbols to provide symmetric privacy as follows

$$\hat{Z}_n(1) = Z'_1(1) + \alpha_n Z'_2(1) + \dots + \alpha_n^{T_1+H-1} Z'_{H+T_1}(1) + \alpha_n^{T_1+H} R'_1(1) + \dots + \alpha_n^{T_1+H+B-1} R'_B(1), \quad (87)$$

$$\hat{Z}_n(2) = Z'_1(2) + \alpha_n Z'_2(2) + \dots + \alpha_n^{T_2+H-1} Z'_{H+T_2}(2) + \alpha_n^{T_2+H} R'_1(2) + \dots + \alpha_n^{T_2+H+B-1} R'_B(2), \quad (88)$$

The queries are given by

$$\mathbf{Q}_n^{[\theta]}(1) = \begin{bmatrix} \frac{1}{f_1 - \alpha_n} \left(\mathbf{e}_\theta + \sum_{j=1}^{T_1} (f_1 - \alpha_n)^j \mathbf{Z}_{1j}(1) \right) \\ \vdots \\ \frac{1}{f_{L_2} - \alpha_n} \left(\mathbf{e}_\theta + \sum_{j=1}^{T_1} (f_{L_2} - \alpha_n)^j \mathbf{Z}_{L_2,j}(1) \right) \end{bmatrix}, \quad (89)$$

and

$$\mathbf{Q}_n^{[\theta]}(2) = \begin{bmatrix} \frac{1}{f_1 - \alpha_n} \left(\mathbf{e}_\theta + \sum_{j=1}^{T_2} (f_1 - \alpha_n)^j \mathbf{Z}_{1j}(2) \right) \\ \vdots \\ \frac{1}{f_{L_2} - \alpha_n} \left(\mathbf{e}_\theta + \sum_{j=1}^{T_2} (f_{L_2} - \alpha_n)^j \mathbf{Z}_{L_2,j}(2) \right) \end{bmatrix}. \quad (90)$$

Each answer instance from the n th server is given by

$$A_n^{[\theta]}(i) = S_n(i)^t \mathbf{Q}_n^{[\theta]}(i) + \hat{Z}_n(i), \quad i \in [2]. \quad (91)$$

Upon applying the same steps as the previous two sections, the rate becomes $R_Q = \frac{L_1+L_2}{N} = \frac{N-6B-2U-E}{N}$.

D. Regime 4: $H + M + B < \frac{N}{2}$ and $2B + U + E > H + M + B$

In this setting, the usage of the classical scheme is preferable since there are not enough dimensions for transmitting message symbols using the classical scheme or the rate provided by the classical scheme is higher than the quantum scheme. For the sake of completeness, we present the classical scheme here. The sub-packetization length is $L = N - H - M - U - 3B$ and the storage can be written as in (27) and queries by the user as in (28). In addition, the answers are computed as

$$A_n^{[\theta]} = S_n^t \mathbf{Q}_n^{[\theta]} + \hat{Z}_n(1), \quad n \in [N] \setminus \mathcal{B} \cup \mathcal{U}. \quad (92)$$

Thus, the answers can be written collectively as

$$\begin{bmatrix} A_{i_1}^{[\theta]} \\ \vdots \\ A_{i_{N-U}}^{[\theta]} \end{bmatrix} = \text{CSA}_{N-U, L, N-U} \mathbf{x}, \quad (93)$$

where

$$\mathbf{x} = \begin{bmatrix} \mathbf{W}_\theta \\ \mathbf{I}_{[H+M]} + \mathbf{Z}'_{[H+M]} \\ \mathbf{R}'_{[B]} \\ \mathbf{0}_{2B} \end{bmatrix} + \text{CSA}_{N-U, L, N-U}^{-1} \Delta_B. \quad (94)$$

Remark 14: Note that in this section we assumed that the Byzantine servers use Pauli $X(\cdot)$, and $Z(\cdot)$ operations when

transmitting the answers. This can be extended in the same exact way that is used in [46]. To explain intuitively, Kraus operators of any quantum operation that output Hilbert space is the same as the input Hilbert space can be decomposed into a linear combination of Weyl operators $W(\nu)$, and any quantum code that corrects errors with Kraus operators $\{E_i\}_i$ can also correct errors with $\{F_i\}$ where $F_i = \sum_j m_{ij} E_i$. For each term individually, the scheme has the same rate thus by using superposition the rate is the same as considering the Weyl operators only being applied by the Byzantine servers.

VII. DECODING THE BYZANTINE ERRORS

In the previous sections, the schemes are analyzed such that the user retrieves the $N - U$ symbols after applying the measurements. However, as discussed earlier, those symbols are contaminated with Byzantine errors. In this section, the decoding technique after measurements is developed in order to perform error correction for the Byzantine errors. We will use the following definitions for ease of notation:

$$\text{CSA}_{N-U, L, N-U}^{-1} = [C_{i,j}], \quad (95)$$

$$\mathbf{D}(a, \mathcal{J}) = \begin{bmatrix} C_{1, \mathcal{J}(1)} & \dots & C_{1, \mathcal{J}(B)} \\ \vdots & \ddots & \vdots \\ C_{a, \mathcal{J}(1)} & \dots & C_{a, \mathcal{J}(B)} \end{bmatrix}, \quad (96)$$

$$\Phi(\mathcal{J}) = \begin{bmatrix} C_{L+X+M+1, \mathcal{J}(1)} & \dots & C_{L+X+M+1, \mathcal{J}(B)} \\ \vdots & \ddots & \vdots \\ C_{L+X+M+B, \mathcal{J}(1)} & \dots & C_{L+X+M+B, \mathcal{J}(B)} \end{bmatrix}, \quad (97)$$

$$\Psi(\mathcal{J}) = \begin{bmatrix} C_{L+X+M+B+1, \mathcal{J}(1)} & \dots & C_{L+X+M+B+1, \mathcal{J}(B)} \\ \vdots & \ddots & \vdots \\ C_{N-U, \mathcal{J}(1)} & \dots & C_{N-U, \mathcal{J}(B)} \end{bmatrix}, \quad (98)$$

$$\hat{\Delta}_B = [\Delta_1, \dots, \Delta_B]^t. \quad (99)$$

Using these definitions, the following lemmas will be useful in the decoding procedure.

Lemma 1: $\Psi(\mathcal{J})$ is invertible for all $\mathcal{J} \subset [N]$, such that $|\mathcal{J}| = B$.

Proof: We prove this lemma by contradiction. Assume for any such $\mathcal{J}$ that $\Psi(\mathcal{J})$ is not invertible. Thus, at least one of the columns can be written as a linear combination of the other columns, i.e.,

$$\begin{bmatrix} C_{L+X+M+B+1, \mathcal{J}(m)} \\ \vdots \\ C_{N-U, \mathcal{J}(m)} \end{bmatrix} = \sum_{\substack{k=1 \\ k \neq m}}^B a_k \begin{bmatrix} C_{L+X+M+B+1, \mathcal{J}(k)} \\ \vdots \\ C_{N-U, \mathcal{J}(k)} \end{bmatrix}. \quad (100)$$

In addition, since $[C_{1, \mathcal{J}(1)}, \dots, C_{N-U, \mathcal{J}(1)}]^t, \dots, [C_{1, \mathcal{J}(B)}, \dots, C_{N-U, \mathcal{J}(B)}]^t$ are orthogonal to exactly $N - U - B$ rows of $\text{CSA}_{N-U, L, N-U}$, we have

$$\begin{bmatrix} \frac{1}{f_1 - \alpha_{\mathcal{I}(1)}} & \dots & \frac{1}{f_L - \alpha_{\mathcal{I}(1)}} & 1 \dots \alpha_{\mathcal{I}(1)}^{X+M+B-1} \\ \frac{1}{f_1 - \alpha_{\mathcal{I}(2)}} & \dots & \frac{1}{f_L - \alpha_{\mathcal{I}(2)}} & 1 \dots \alpha_{\mathcal{I}(2)}^{X+M+B-1} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{f_1 - \alpha_{\mathcal{I}(N-U-B)}} & \dots & \frac{1}{f_L - \alpha_{\mathcal{I}(N-U-B)}} & 1 \dots \alpha_{\mathcal{I}(N-U-B)}^{X+M+B-1} \end{bmatrix} \times \begin{bmatrix} C_{1, \mathcal{J}(1)} & \dots & C_{1, \mathcal{J}(B)} \\ \vdots & \vdots & \vdots \\ C_{N-U, \mathcal{J}(1)} & \dots & C_{N-U, \mathcal{J}(B)} \end{bmatrix} = 0, \quad (101)$$

where $\mathcal{I} \cup \mathcal{J} = [N]$, and $\mathcal{I} \cap \mathcal{J} = \emptyset$. In addition, we have

$$\begin{aligned}
 & \begin{bmatrix} \frac{1}{f_1 - \alpha_{\mathcal{I}(1)}} & \cdots & \frac{1}{f_L - \alpha_{\mathcal{I}(1)}} & 1 \cdot \alpha_{\mathcal{I}(1)}^{X+M+B-1} \\ \frac{1}{f_1 - \alpha_{\mathcal{I}(2)}} & \cdots & \frac{1}{f_L - \alpha_{\mathcal{I}(2)}} & 1 \cdot \alpha_{\mathcal{I}(2)}^{X+M+B-1} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{f_1 - \alpha_{\mathcal{I}(N-U-B)}} & \cdots & \frac{1}{f_L - \alpha_{\mathcal{I}(N-U-B)}} & 1 \cdot \alpha_{\mathcal{I}(N-U-B)}^{X+M+B-1} \end{bmatrix} \\
 & \times \begin{bmatrix} C_{1,\mathcal{J}(1)} & \cdots & C_{1,\mathcal{J}(B)} \\ \vdots & \vdots & \vdots \\ C_{L+X+M+B,\mathcal{J}(1)} & \cdots & C_{L+X+M+B,\mathcal{J}(B)} \end{bmatrix} \\
 & = - \begin{bmatrix} \alpha_{\mathcal{I}(1)}^{X+M+B} & \cdots & \alpha_{\mathcal{I}(1)}^{N-U} \\ \alpha_{\mathcal{I}(2)}^{X+M+B} & \cdots & \alpha_{\mathcal{I}(2)}^{N-U} \\ \vdots & \vdots & \vdots \\ \alpha_{\mathcal{I}(N-U-B)}^{X+M+B} & \cdots & \alpha_{\mathcal{I}(N-U-B)}^{N-U} \end{bmatrix} \\
 & \times \begin{bmatrix} C_{L+X+M+B+1,\mathcal{J}(1)} & \cdots & C_{L+X+M+B+1,\mathcal{J}(B)} \\ \vdots & \vdots & \vdots \\ C_{N-U,\mathcal{J}(1)} & \cdots & C_{N-U,\mathcal{J}(B)} \end{bmatrix}, \quad (102)
 \end{aligned}$$

thus,

$$\begin{aligned}
 & \begin{bmatrix} \frac{1}{f_1 - \alpha_{\mathcal{I}(1)}} & \cdots & \frac{1}{f_L - \alpha_{\mathcal{I}(1)}} & 1 \cdot \alpha_{\mathcal{I}(1)}^{X+M+B-1} \\ \frac{1}{f_1 - \alpha_{\mathcal{I}(2)}} & \cdots & \frac{1}{f_L - \alpha_{\mathcal{I}(2)}} & 1 \cdot \alpha_{\mathcal{I}(2)}^{X+M+B-1} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{f_1 - \alpha_{\mathcal{I}(N-U-B)}} & \cdots & \frac{1}{f_L - \alpha_{\mathcal{I}(N-U-B)}} & 1 \cdot \alpha_{\mathcal{I}(N-U-B)}^{X+M+B-1} \end{bmatrix} \begin{bmatrix} C_{1,\mathcal{J}(m)} \\ \vdots \\ C_{L+X+M+B,\mathcal{J}(m)} \end{bmatrix} \\
 & = - \begin{bmatrix} \alpha_{\mathcal{I}(1)}^{X+M+B} & \cdots & \alpha_{\mathcal{I}(1)}^{N-U} \\ \alpha_{\mathcal{I}(2)}^{X+M+B} & \cdots & \alpha_{\mathcal{I}(2)}^{N-U} \\ \vdots & \vdots & \vdots \\ \alpha_{\mathcal{I}(N-U-B)}^{X+M+B} & \cdots & \alpha_{\mathcal{I}(N-U-B)}^{N-U} \end{bmatrix} \begin{bmatrix} C_{L+X+M+B+1,\mathcal{J}(m)} \\ \vdots \\ C_{N-U,\mathcal{J}(m)} \end{bmatrix} \quad (103) \\
 & = - \sum_{\substack{k=1 \\ k \neq m}} a_k \begin{bmatrix} \alpha_{\mathcal{I}(1)}^{X+M+B} & \cdots & \alpha_{\mathcal{I}(1)}^{N-U} \\ \alpha_{\mathcal{I}(2)}^{X+M+B} & \cdots & \alpha_{\mathcal{I}(2)}^{N-U} \\ \vdots & \vdots & \vdots \\ \alpha_{\mathcal{I}(N-U-B)}^{X+M+B} & \cdots & \alpha_{\mathcal{I}(N-U-B)}^{N-U} \end{bmatrix} \begin{bmatrix} C_{L+X+M+B+1,\mathcal{J}(k)} \\ \vdots \\ C_{N-U,\mathcal{J}(k)} \end{bmatrix} \quad (104)
 \end{aligned}$$

$$\begin{aligned}
 & = \sum_{\substack{k=1 \\ k \neq m}} a_k \begin{bmatrix} \frac{1}{f_1 - \alpha_{\mathcal{I}(1)}} & \cdots & \frac{1}{f_L - \alpha_{\mathcal{I}(1)}} & 1 \cdot \alpha_{\mathcal{I}(1)}^{X+M+B-1} \\ \frac{1}{f_1 - \alpha_{\mathcal{I}(2)}} & \cdots & \frac{1}{f_L - \alpha_{\mathcal{I}(2)}} & 1 \cdot \alpha_{\mathcal{I}(2)}^{X+M+B-1} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{f_1 - \alpha_{\mathcal{I}(N-U-B)}} & \cdots & \frac{1}{f_L - \alpha_{\mathcal{I}(N-U-B)}} & 1 \cdot \alpha_{\mathcal{I}(N-U-B)}^{X+M+B-1} \end{bmatrix} \\
 & \times \begin{bmatrix} C_{1,\mathcal{J}(k)} \\ \vdots \\ C_{L+X+M+B,\mathcal{J}(k)} \end{bmatrix}. \quad (105)
 \end{aligned}$$

This means that the first $L + X + M + B$ components of the m th column vector is as well a linear combination of the other columns with the same coefficients, i.e.,

$$\begin{bmatrix} C_{1,\mathcal{J}(m)} \\ \vdots \\ C_{N-U,\mathcal{J}(m)} \end{bmatrix} = \sum_{\substack{k=1 \\ k \neq m}} a_k \begin{bmatrix} C_{1,\mathcal{J}(k)} \\ \vdots \\ C_{N-U,\mathcal{J}(k)} \end{bmatrix}. \quad (106)$$

This is a contradiction since the inverse matrix of $\text{CSA}_{N-U,L,N-U}$ is a full-rank matrix. ■

Lemma 2: If $\Phi(\mathcal{J}_1)\Psi(\mathcal{J}_1)^{-1}\Psi(\mathcal{J}_2)\hat{\Delta}_B = \Phi(\mathcal{J}_2)\hat{\Delta}_B$, then $D(N - U - 2B, \mathcal{J}_1)\Psi(\mathcal{J}_1)^{-1}\Psi(\mathcal{J}_2)\hat{\Delta}_B = D(N - U - 2B, \mathcal{J}_2)\hat{\Delta}_B$, and $D(N - U, \mathcal{J}_1)\Psi(\mathcal{J}_1)^{-1}\Psi(\mathcal{J}_2)\hat{\Delta}_B = D(N - U, \mathcal{J}_2)\hat{\Delta}_B$.

Proof: First, for any $\mathcal{J}_1$, and $\mathcal{J}_2$, we have $|\mathcal{J}_1 \cup \mathcal{J}_2| \leq 2B$. Thus, there are at least $2B$ rows of $\text{CSA}_{N-U,L,N-U}$ that are orthogonal to $C_{:,m}$, such that $m \in \mathcal{J}_1 \cup \mathcal{J}_2$. Thus, as in the previous proof, we have

$$\begin{aligned}
 & \begin{bmatrix} \frac{1}{f_1 - \alpha_{\mathcal{I}(1)}} & \cdots & \frac{1}{f_L - \alpha_{\mathcal{I}(1)}} & 1 \cdot \alpha_{\mathcal{I}(1)}^{X+M-1} \\ \frac{1}{f_1 - \alpha_{\mathcal{I}(2)}} & \cdots & \frac{1}{f_L - \alpha_{\mathcal{I}(2)}} & 1 \cdot \alpha_{\mathcal{I}(2)}^{X+M-1} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{f_1 - \alpha_{\mathcal{I}(N-U-2B)}} & \cdots & \frac{1}{f_L - \alpha_{\mathcal{I}(N-U-2B)}} & 1 \cdot \alpha_{\mathcal{I}(N-U-2B)}^{X+M-1} \end{bmatrix} \\
 & \times \begin{bmatrix} C_{1,\mathcal{J}(1)} & \cdots & C_{1,\mathcal{J}(B)} \\ \vdots & \vdots & \vdots \\ C_{L+X+M,\mathcal{J}(1)} & \cdots & C_{L+X+M,\mathcal{J}(B)} \end{bmatrix} \\
 & = - \begin{bmatrix} \alpha_{\mathcal{I}(1)}^{X+M} & \cdots & \alpha_{\mathcal{I}(1)}^{N-U} \\ \alpha_{\mathcal{I}(2)}^{X+M} & \cdots & \alpha_{\mathcal{I}(2)}^{N-U} \\ \vdots & \vdots & \vdots \\ \alpha_{\mathcal{I}(N-U-2B)}^{X+M} & \cdots & \alpha_{\mathcal{I}(N-U-2B)}^{N-U} \end{bmatrix} \\
 & \times \begin{bmatrix} C_{L+X+M+1,\mathcal{J}(1)} & \cdots & C_{L+X+M+1,\mathcal{J}(B)} \\ \vdots & \vdots & \vdots \\ C_{N-U,\mathcal{J}(1)} & \cdots & C_{N-U,\mathcal{J}(B)} \end{bmatrix}, \quad (107)
 \end{aligned}$$

which can be written as

$$\begin{aligned}
 & \begin{bmatrix} \frac{1}{f_1 - \alpha_{\mathcal{I}(1)}} & \cdots & \frac{1}{f_L - \alpha_{\mathcal{I}(1)}} & 1 \cdot \alpha_{\mathcal{I}(1)}^{X+M-1} \\ \frac{1}{f_1 - \alpha_{\mathcal{I}(2)}} & \cdots & \frac{1}{f_L - \alpha_{\mathcal{I}(2)}} & 1 \cdot \alpha_{\mathcal{I}(2)}^{X+M-1} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{f_1 - \alpha_{\mathcal{I}(N-U-B)}} & \cdots & \frac{1}{f_L - \alpha_{\mathcal{I}(N-U-B)}} & 1 \cdot \alpha_{\mathcal{I}(N-U-B)}^{X+M-1} \end{bmatrix} \\
 & \times D(N - U - 2B, \mathcal{J}_1) \\
 & = - \begin{bmatrix} \alpha_{\mathcal{I}(1)}^{X+M} & \cdots & \alpha_{\mathcal{I}(1)}^{N-U} \\ \alpha_{\mathcal{I}(2)}^{X+M} & \cdots & \alpha_{\mathcal{I}(2)}^{N-U} \\ \vdots & \vdots & \vdots \\ \alpha_{\mathcal{I}(N-U-2B)}^{X+M} & \cdots & \alpha_{\mathcal{I}(N-U-2B)}^{N-U} \end{bmatrix} \begin{bmatrix} \Phi(\mathcal{J}_1) \\ \Psi(\mathcal{J}_1) \end{bmatrix}. \quad (108)
 \end{aligned}$$

By multiplying both sides by $\Psi(\mathcal{J}_1)^{-1}\Psi(\mathcal{J}_2)\hat{\Delta}_B$, we have

$$\begin{aligned}
 & \begin{bmatrix} \frac{1}{f_1 - \alpha_{\mathcal{I}(1)}} & \cdots & \frac{1}{f_L - \alpha_{\mathcal{I}(1)}} & 1 \cdot \alpha_{\mathcal{I}(1)}^{X+M-1} \\ \frac{1}{f_1 - \alpha_{\mathcal{I}(2)}} & \cdots & \frac{1}{f_L - \alpha_{\mathcal{I}(2)}} & 1 \cdot \alpha_{\mathcal{I}(2)}^{X+M-1} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{f_1 - \alpha_{\mathcal{I}(N-U-B)}} & \cdots & \frac{1}{f_L - \alpha_{\mathcal{I}(N-U-B)}} & 1 \cdot \alpha_{\mathcal{I}(N-U-B)}^{X+M-1} \end{bmatrix} \\
 & \times D(N - U - 2B, \mathcal{J}_1)\Psi(\mathcal{J}_1)^{-1}\Psi(\mathcal{J}_2)\hat{\Delta}_B \\
 & = - \begin{bmatrix} \alpha_{\mathcal{I}(1)}^{X+M} & \cdots & \alpha_{\mathcal{I}(1)}^{N-U} \\ \alpha_{\mathcal{I}(2)}^{X+M} & \cdots & \alpha_{\mathcal{I}(2)}^{N-U} \\ \vdots & \vdots & \vdots \\ \alpha_{\mathcal{I}(N-U-2B)}^{X+M} & \cdots & \alpha_{\mathcal{I}(N-U-2B)}^{N-U} \end{bmatrix} \\
 & \times \begin{bmatrix} \Phi(\mathcal{J}_1)\Psi(\mathcal{J}_1)^{-1}\Psi(\mathcal{J}_2)\hat{\Delta}_B \\ \Psi(\mathcal{J}_2)\hat{\Delta}_B \end{bmatrix} \quad (109)
 \end{aligned}$$

$$= - \begin{bmatrix} \alpha_{\mathcal{I}(1)}^{X+M} & \cdots & \alpha_{\mathcal{I}(1)}^{N-U} \\ \alpha_{\mathcal{I}(2)}^{X+M} & \cdots & \alpha_{\mathcal{I}(2)}^{N-U} \\ \vdots & \vdots & \vdots \\ \alpha_{\mathcal{I}(N-U-2B)}^{X+M} & \cdots & \alpha_{\mathcal{I}(N-U-2B)}^{N-U} \end{bmatrix} \begin{bmatrix} \Phi(\mathcal{J}_2) \hat{\Delta}_B \\ \Psi(\mathcal{J}_2) \hat{\Delta}_B \end{bmatrix} \quad (110)$$

$$= \begin{bmatrix} \frac{1}{f_1 - \alpha_{\mathcal{I}(1)}} & \cdots & \frac{1}{f_L - \alpha_{\mathcal{I}(1)}} & 1 \cdots \alpha_{\mathcal{I}(1)}^{X+M-1} \\ \frac{1}{f_1 - \alpha_{\mathcal{I}(2)}} & \cdots & \frac{1}{f_L - \alpha_{\mathcal{I}(2)}} & 1 \cdots \alpha_{\mathcal{I}(2)}^{X+M-1} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{f_1 - \alpha_{\mathcal{I}(N-U-2B)}} & \cdots & \frac{1}{f_L - \alpha_{\mathcal{I}(N-U-2B)}} & 1 \cdots \alpha_{\mathcal{I}(N-U-2B)}^{X+M-1} \end{bmatrix} \times D(N-U-2B, \mathcal{J}_2) \hat{\Delta}_B. \quad (111)$$

This concludes the first part of the lemma. The proof of the second part follows as

$$D(N-U, \mathcal{J}_2) \hat{\Delta}_B = \begin{bmatrix} V(N-U-2B, \mathcal{J}_2) \\ \Phi(\mathcal{J}_1) \\ \Psi(\mathcal{J}_1) \end{bmatrix} \hat{\Delta}_B \quad (112)$$

$$= \begin{bmatrix} D(N-U-2B, \mathcal{J}_1) \Psi(\mathcal{J}_1)^{-1} \Psi(\mathcal{J}_2) \hat{\Delta}_B \\ \Phi(\mathcal{J}_1) \Psi(\mathcal{J}_1)^{-1} \Psi(\mathcal{J}_2) \hat{\Delta}_B \\ \Psi(\mathcal{J}_1) \Psi(\mathcal{J}_1)^{-1} \Psi(\mathcal{J}_2) \hat{\Delta}_B \end{bmatrix} \quad (113)$$

$$= \begin{bmatrix} D(N-U-2B, \mathcal{J}_1) \\ \Phi(\mathcal{J}_1) \\ \Psi(\mathcal{J}_1) \end{bmatrix} \Psi(\mathcal{J}_1)^{-1} \Psi(\mathcal{J}_2) \hat{\Delta}_B \quad (114)$$

$$= D(N-U, \mathcal{J}_1) \Psi(\mathcal{J}_1)^{-1} \Psi(\mathcal{J}_2) \hat{\Delta}_B, \quad (115)$$

concluding the proof. ■

Now, assume that the real Byzantine servers are $\mathcal{J}_2$ and the user guess is $\mathcal{J}_1$. Since we have $2B$ symbols, we can do the following: First, the user estimates $\hat{\Delta}_B$ using the first B symbols and using the last B symbols, the estimation is either confirmed or denied. Since $\Psi(\mathcal{J}_1)$ is invertible, the user gets the following estimate $\bar{\Delta}_B = \Psi(\mathcal{J}_1)^{-1} \Psi(\mathcal{J}_2) \hat{\Delta}_B$. Afterwards, the user checks if $\Phi(\mathcal{J}_1) \Psi(\mathcal{J}_1)^{-1} \Psi(\mathcal{J}_2) \hat{\Delta}_B = \Phi(\mathcal{J}_2) \hat{\Delta}_B$. If this is satisfied, then to remove the Byzantine effect from the $N-U$ servers, the user multiplies $\bar{\Delta}_B$ by $D(N-U, \mathcal{J}_1)$ and,

$$\begin{aligned} & \begin{bmatrix} W_\theta(i) \\ I_{[X+M-\gamma_1]}(i) + Z'_{[X+M-\gamma_1]}(i) \end{bmatrix} \\ &= \mathbf{x}(1 : L + X + M - \gamma_1) \\ & \quad - (D(N-U, \mathcal{J}_1) \bar{\Delta}_B) (1 : L + X + M - \gamma_1), \end{aligned} \quad (116)$$

where $\gamma_1 = \lfloor \frac{N}{2} \rfloor$, and $\gamma_2 = \lceil \frac{N}{2} \rceil$.

VIII. PROOFS

Lemma 3: The storage at the n th server given by

$$S_n = \begin{bmatrix} W_{\cdot,1} + \sum_{j=1}^X (f_1 - \alpha_n)^j R_{1j} \\ W_{\cdot,2} + \sum_{j=1}^X (f_2 - \alpha_n)^j R_{2j} \\ \vdots \\ W_{\cdot,L} + \sum_{j=1}^X (f_L - \alpha_n)^j R_{Lj} \end{bmatrix}, \quad (117)$$

is secure against any X communicating databases.

Proof: First, note that since each symbol and each message are encoded using different random variables, it suffices to prove security for only one symbol. Let the ℓ th symbol of the

m th message stored in the n th database be denoted by $\tilde{W}_{m,\ell}(n)$, thus

$$\tilde{W}_{m,\ell}(n) = W_{m,\ell} + (f_\ell - \alpha_n) r_1 + \dots + (f_\ell - \alpha_n)^X r_X, \quad (118)$$

where r_i , $i \in [X]$ are i.i.d. random variables generated uniformly at random. After sharing $\tilde{W}_{m,\ell}(n_1), \dots, \tilde{W}_{m,\ell}(n_X)$, each server has

$$W_{\mathcal{X}} = \begin{bmatrix} W_{m,\ell} + (f_\ell - \alpha_{n_1}) r_1 + \dots + (f_\ell - \alpha_{n_1})^X r_X \\ \vdots \\ W_{m,\ell} + (f_\ell - \alpha_{n_X}) r_1 + \dots + (f_\ell - \alpha_{n_X})^X r_X \end{bmatrix} \quad (119)$$

$$= \mathbf{1} W_{m,\ell} + \underbrace{\begin{bmatrix} B_1 \\ \vdots \\ B_X \end{bmatrix} \begin{bmatrix} r_1 \\ \vdots \\ r_X \end{bmatrix}}_{\mathbf{r}}, \quad (120)$$

where $B_i = [f_\ell - \alpha_{n_i}, \dots, (f_\ell - \alpha_{n_i})^X]$. Since each row B_ℓ is independent, we know that $I(B_\ell \mathbf{r}; B_{\ell'} \mathbf{r}) = 0$, and $I(B_\ell \mathbf{r}; B_{\ell'} \mathbf{r} | \{B_k \mathbf{r}\}_{k \in [X] \setminus \{\ell, \ell'\}}) = 0$ using Shannon's one-time pad theorem. Thus,

$$I(W_{m,\ell}; W_{\mathcal{X}}) = \sum_{k \in [X]} I(W_{m,\ell}; W_{m,\ell} + B_k \mathbf{r}) \quad (121)$$

$$= \sum_{k \in [X]} I(W_{m,\ell}; B_k \mathbf{r}) = 0, \quad (122)$$

completing the proof. ■

Lemma 4: The queries transmitted to the n th server given by

$$\mathbf{Q}_n^{[\theta]} = \begin{bmatrix} \frac{1}{f_1 - \alpha_n} \left(\mathbf{e}_\theta + \sum_{j=1}^M (f_1 - \alpha_n)^j \mathbf{Z}_{1j} \right) \\ \vdots \\ \frac{1}{f_L - \alpha_n} \left(\mathbf{e}_\theta + \sum_{j=1}^M (f_L - \alpha_n)^j \mathbf{Z}_{Lj} \right) \end{bmatrix} = \begin{bmatrix} \mathbf{q}_{n,1} \\ \vdots \\ \mathbf{q}_{n,L} \end{bmatrix}, \quad (123)$$

are private against any $T \leq M$ colluding servers.

Proof: First, assume that $T = M$ which is the largest possible value for T . Now, since each $\mathbf{q}_{n,i}$, $i \in [L]$ is independent of all others (as the random vectors used at each query instance are independent) it is sufficient to consider $\mathbf{q}_{T,i} = [\mathbf{q}_{n_1,i}, \dots, \mathbf{q}_{n_M,i}]$. This can be written as

$$\begin{aligned} \begin{bmatrix} \mathbf{q}_{n_1,1} \\ \vdots \\ \mathbf{q}_{n_M,1} \end{bmatrix} &= \begin{bmatrix} \frac{1}{f_1 - \alpha_{n_1}} \\ \vdots \\ \frac{1}{f_1 - \alpha_{n_M}} \end{bmatrix} \mathbf{e}_\theta \\ &+ \underbrace{\begin{bmatrix} 1 & (f_1 - \alpha_{n_1}) & \cdots & (f_1 - \alpha_{n_1})^{M-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & (f_1 - \alpha_{n_M}) & \cdots & (f_1 - \alpha_{n_M})^{M-1} \end{bmatrix}}_{\mathbf{B}} \underbrace{\begin{bmatrix} \mathbf{Z}_{11} \\ \vdots \\ \mathbf{Z}_{1M} \end{bmatrix}}_{\mathbf{Z}}. \end{aligned} \quad (124)$$

As explained in the previous proof, the rows of $\mathbf{B}$ are independent from each other and with the same line of arguments, we reach the required conclusion. ■

Remark 15: Note that Lemma 4 implies $I(\theta; \mathbf{Q}_{\mathcal{E}}^{[\theta]}) = 0$, for $|\mathcal{E}| \leq M$.

In the next part, we prove some properties related to the transmitted qudits and measurements. Although this can be proven using the Von Neumann entropy and quantum mutual

information as in [7], another direction is used here using the classical information theoretic quantities in order to provide readers with an alternative way to deal with quantum systems using the N -sum box abstraction. We prove the required properties for the scheme for Byzantine servers with dynamic eavesdroppers, since it is the general version. Recall that for the Byzantine scheme presented in Section VI, the transmitted vector for the first regime is written as

$$\mathbf{h}_1 = \begin{bmatrix} \mathbf{I}_N & 0 & 0 \\ 0 & \tilde{\mathbf{V}}_4 & 0 \\ 0 & 0 & \mathbf{I}_{4B+2U} \end{bmatrix} \underbrace{\begin{bmatrix} \tilde{\mathbf{I}}_1(1) \\ \tilde{\mathbf{I}}_1(2) \\ \tilde{\mathbf{W}}_\theta(1) \\ \tilde{\mathbf{W}}_\theta(2) \\ \tilde{\mathbf{I}}_2(1) \\ \tilde{\mathbf{I}}_2(2) \\ \tilde{\Delta}_B(1) \\ \tilde{\Delta}_B(2) \\ \Delta'_U(1) \\ \Delta'_U(2) \end{bmatrix}}_x, \quad (125)$$

where $\tilde{\mathbf{V}}_4 = \mathbf{V}_{2L+2H+2M-N}(\mathbf{b}_{[2L+2H+2M+2B-N]})$,

$$\begin{bmatrix} \tilde{\mathbf{W}}_\theta(i) \\ \tilde{\mathbf{I}}_1(i) \\ \tilde{\mathbf{I}}_2(i) \\ \tilde{\Delta}_B(i) \end{bmatrix} = \begin{bmatrix} \mathbf{W}_\theta(i) \\ \mathbf{I}_{[H+M+B]}(i) \\ \mathbf{0}_{2B} \end{bmatrix} + \mathbf{CSA}_{N-U,L,N-U}^{-1} \Delta_B(i), \quad (126)$$

and

$$\hat{\mathbf{I}}_{[H+M+B]}(i) = \begin{bmatrix} \mathbf{I}_{[H+M]}(i) + \mathbf{Z}'_{[H+M]}(i) \\ \mathbf{R}'_{[B]}(i) \end{bmatrix}. \quad (127)$$

Upon measurement, the user retrieves N symbols of $\mathbf{h}_1$, counting the erasure errors as well. In this paper, we examine the vector $\mathbf{h}_1$ and show that symmetric privacy and eavesdropper security are maintained for whatever $N - U$ or $E \leq N - U$ linear combination that can come out of any measurements.

Lemma 5: The noise symbols used in masking are secure against any B Byzantine servers.

Proof: Let $\mathcal{B} = \{i_1, \dots, i_B\}$ denote the indices of the Byzantine servers. If the B Byzantine servers are cooperating, they will have the noise that should have been added to the generated answers if they were honest as

$$\begin{aligned} \begin{bmatrix} \hat{Z}_{i_1} \\ \vdots \\ \hat{Z}_{i_B} \end{bmatrix} &= \begin{bmatrix} 1 & \alpha_{i_1} & \dots & \alpha_{i_1}^{H+M+B-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_{i_B} & \dots & \alpha_{i_B}^{H+M+B-1} \end{bmatrix} \begin{bmatrix} Z'_1 \\ \vdots \\ Z'_{H+M} \\ R'_1 \\ \vdots \\ R'_B \end{bmatrix} \\ &= \begin{bmatrix} 1 & \dots & \alpha_{i_1}^{H+M-1} \\ \vdots & \ddots & \vdots \\ 1 & \dots & \alpha_{i_B}^{H+M-1} \end{bmatrix} \begin{bmatrix} Z'_1 \\ \vdots \\ Z'_{H+M} \end{bmatrix} \\ &\quad + \begin{bmatrix} \alpha_{i_1}^{H+M} & \dots & \alpha_{i_1}^{H+M+B-1} \\ \vdots & \ddots & \vdots \\ \alpha_{i_B}^{H+M} & \dots & \alpha_{i_B}^{H+M+B-1} \end{bmatrix} \begin{bmatrix} R'_1 \\ \vdots \\ R'_B \end{bmatrix}, \quad (129) \end{aligned}$$

for each instance of the answers, however, as $Z'_1(1), \dots, Z'_{H+M}(1), R'_1(1), \dots, R'_B(1)$ are independent of

$Z'_1(2), \dots, Z'_{H+M}(2), R'_1(2), \dots, R'_B(2)$, it suffices to focus on only one instance to show the required result, as the other instance will follow from the same arguments.

For brevity, let

$$\boldsymbol{\alpha}_B = \begin{bmatrix} \alpha_{i_1} \\ \vdots \\ \alpha_{i_B} \end{bmatrix}. \quad (130)$$

Moreover, let $\mathbf{V}(\boldsymbol{\alpha}_B)_{B \times H+M}$ denote the Vandermonde matrix of $B \times H + M$ dimensions and $\boldsymbol{\alpha}_B$ as the interpolation points. Thus, (129) in this notation is written as

$$\hat{\mathbf{Z}}_B = \mathbf{V}(\boldsymbol{\alpha}_B)_{B \times H+M} \mathbf{Z}'_{[H+M]} + \text{diag}(\boldsymbol{\alpha}_B)^{H+M} \mathbf{V}(\boldsymbol{\alpha}_B)_{B \times B} \mathbf{R}'_{[B]}, \quad (131)$$

which means that

$$\mathbb{P}(\hat{\mathbf{Z}}_B = \hat{\mathbf{z}}) = \frac{1}{q^B} \frac{1}{q^{H+M}} q^{H+M} = \frac{1}{q^B}, \quad (132)$$

where it follows from the fact that, for any value of $\mathbf{R}'_{[B]}$, there exists a $\mathbf{Z}'_{[H+M]}$ satisfying the event inside the probability. This showcases that $\hat{\mathbf{Z}}_B$ is uniformly distributed.

Now, also note that

$$\begin{aligned} \mathbb{P}(\hat{\mathbf{Z}}_B = \hat{\mathbf{z}}, \mathbf{Z}'_{[H+M]} = \mathbf{z}') &= \mathbb{P}(\mathbf{V}(\boldsymbol{\alpha}_B)_{B \times H+M} \mathbf{Z}'_{[H+M]} \\ &\quad + \text{diag}(\boldsymbol{\alpha}_B)^{H+M} \mathbf{V}(\boldsymbol{\alpha}_B)_{B \times B} \mathbf{R}'_{[B]} = \hat{\mathbf{z}}, \mathbf{Z}'_{[H+M]} = \mathbf{z}') \end{aligned} \quad (133)$$

$$= \mathbb{P}(\text{diag}(\boldsymbol{\alpha}_B)^{H+M} \mathbf{V}(\boldsymbol{\alpha}_B)_{B \times B} \mathbf{R}'_{[B]} = \hat{\mathbf{z}} - \mathbf{V}(\boldsymbol{\alpha}_B)_{B \times H+M} \mathbf{z}', \mathbf{Z}'_{[H+M]} = \mathbf{z}') \quad (134)$$

$$= \mathbb{P}(\mathbf{R}'_{[B]} = (\text{diag}(\boldsymbol{\alpha}_B)^{H+M} \mathbf{V}(\boldsymbol{\alpha}_B)_{B \times B})^{-1} \times (\hat{\mathbf{z}} - \mathbf{V}(\boldsymbol{\alpha}_B)_{B \times H+M} \mathbf{z}'), \mathbf{Z}'_{[H+M]} = \mathbf{z}') \quad (135)$$

$$= \mathbb{P}(\mathbf{R}'_{[B]} = (\text{diag}(\boldsymbol{\alpha}_B)^{H+M} \mathbf{V}(\boldsymbol{\alpha}_B)_{B \times B})^{-1} \times (\hat{\mathbf{z}} - \mathbf{V}(\boldsymbol{\alpha}_B)_{B \times H+M} \mathbf{z}')) \mathbb{P}(\mathbf{Z}'_{[H+M]} = \mathbf{z}') \quad (136)$$

$$= \frac{1}{q^B} \frac{1}{q^{H+M}} \quad (137)$$

$$= \mathbb{P}(\hat{\mathbf{Z}}_B = \hat{\mathbf{z}}) \mathbb{P}(\mathbf{Z}'_{[H+M]} = \mathbf{z}'), \quad (138)$$

which shows that $\hat{\mathbf{Z}}$ is independent from $\mathbf{Z}'_{[H+M]}$. Moreover, by construction $Z'_1, \dots, Z'_{H+M}$ are independent from the queries as well. Thus,

$$I(\hat{Z}_{i_1}, \dots, \hat{Z}_{i_B}, \mathbf{Q}_{i_1}^{[\theta]}, \dots, \mathbf{Q}_{i_B}^{[\theta]}, Z'_1, \dots, Z'_{H+M}) = 0, \quad (139)$$

concluding the proof. ■

Lemma 6: The noise symbols used in protecting the interference symbols are secure against the user for any action performed by the B Byzantine servers.

Proof: First, note that $B \leq \frac{N}{3}$. Let $\mathcal{L}_1, \mathcal{L}_2 \subset [H + M]$ with $|\mathcal{L}_1| = |\mathcal{L}_2| = \max\{0, H + M - \lfloor \frac{N}{2} \rfloor\}$ and $\mathcal{H}_1 = \mathcal{H}_2 = \{\max\{1, \lfloor \frac{N}{2} \rfloor - H - M + 1\}, \dots, B\}$. Then, we proceed as follows

$$\begin{aligned} &\frac{1}{2} I(Z'_{\mathcal{L}_1}(1), Z'_{\mathcal{L}_2}(2); \mathbf{R}'_{\mathcal{H}_1}(1), \mathbf{R}'_{\mathcal{H}_2}(2), \hat{\mathbf{Z}}_B(1), \hat{\mathbf{Z}}_B(2)) \\ &= I(Z'_{\mathcal{L}_1}(1); \mathbf{R}'_{\mathcal{H}_1}(1), \hat{\mathbf{Z}}_B(1)) \end{aligned} \quad (140)$$

$$\begin{aligned} &= \underbrace{I(\mathbf{R}'_{\mathcal{H}_1}(1); Z'_{\mathcal{L}_1}(1))}_{(a)} \\ &\quad + \mathbb{1}\left(H + M > \lfloor \frac{N}{2} \rfloor\right) I(Z'_{\mathcal{L}_1}(1); \hat{\mathbf{Z}}_B(1) | \mathbf{R}'_{\mathcal{H}_1}(1)) \end{aligned} \quad (141)$$

$$\begin{aligned}
 &= \mathbb{1} \left(H + M > \lfloor \frac{N}{2} \rfloor \right) \\
 &\times I \left(\mathbf{Z}'_{\mathcal{L}_1}(1); \begin{bmatrix} 1 & \alpha_{n_1} & \dots & \alpha_{n_1}^{H+M-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_{n_B} & \dots & \alpha_{n_B}^{H+M-1} \end{bmatrix} \begin{bmatrix} \mathbf{Z}'_1 \\ \vdots \\ \mathbf{Z}'_{H+M} \end{bmatrix} \right) \\
 &= 0,
 \end{aligned} \tag{142}$$

$$\tag{143}$$

where $(a) = 0$ since these terms are generated uniformly at random and independent from each other, and (143) follows from the fact that $B < \frac{N}{2}$. ■

Remark 16: To make it easier to understand intuitively the proof of Lemma 6: $\mathcal{L}_1$ and $\mathcal{L}_2$ are the interference terms that are not dropped over the air, i.e., decoded by the user upon measurement. In addition, $\mathcal{H}_1$ and $\mathcal{H}_2$ are the remaining masking terms for the shared randomness between the servers. They are all contaminated by the Byzantine noise, however, since the user decodes the Byzantine noise as explained in Section VII, we drop the noise effect. As a numerical example, if $N = 17$, $X = 5$, $T = 4$, $B = 2$, and $E = U = 0$, then $\mathcal{L}_1 = \emptyset$, and $\mathcal{L}_2 = \{9\}$, i.e., no interference terms are received from the first instance while one interference term from the second instance is decoded by the user, $I_9(2)$. On the other hand $\mathcal{H}_1 = \mathcal{H}_2 = \{1, 2\}$, thus $R_1(i)$, $R_2(i)$, for $i \in [2]$ are received.

Lemma 7: For the first regime in Theorem 3, symmetric privacy is guaranteed.

Proof: We proceed as follows,

$$\begin{aligned}
 &I(\mathbf{W}_{[K] \setminus \{\theta\}}; (\mathbf{M}'_1 \mathbf{x})_{[N+1, 2N-2U]}) \\
 &= I(\mathbf{W}_{[K] \setminus \{\theta\}}; (\mathbf{x})_{[N+1, 2N-2U]})
 \end{aligned} \tag{144}$$

$$\begin{aligned}
 &= I(\mathbf{W}_{[K] \setminus \{\theta\}}; \tilde{\Delta}_B(1), \tilde{\Delta}_B(2)) \\
 &\quad + I(\mathbf{W}_{[K] \setminus \{\theta\}}; (\mathbf{x})_{[N+1, 2N-2U-4B]} | \tilde{\Delta}_B(1), \tilde{\Delta}_B(2))
 \end{aligned} \tag{145}$$

$$\begin{aligned}
 &= \underbrace{I(\mathbf{W}_{[K] \setminus \{\theta\}}; \delta_1(1), \dots, \delta_B(2))}_{(a)} \\
 &\quad + I(\mathbf{W}_{[K] \setminus \{\theta\}}; (\mathbf{x})_{[N+1, 2N-2U-4B]} | \delta_1(1), \dots, \delta_B(2))
 \end{aligned} \tag{146}$$

$$\begin{aligned}
 &= I(\mathbf{W}_{[K] \setminus \{\theta\}}; \mathbf{W}_\theta(1), \mathbf{W}_\theta(2), \hat{\mathbf{I}}_2(1), \hat{\mathbf{I}}_2(2) | \delta_1(1), \dots, \delta_B(2)) \\
 &\tag{147}
 \end{aligned}$$

$$\begin{aligned}
 &= I(\mathbf{W}_{[K] \setminus \{\theta\}}; \hat{\mathbf{I}}_2(1), \hat{\mathbf{I}}_2(2) | \delta_1(1), \dots, \delta_B(2), \mathbf{W}_\theta(1), \mathbf{W}_\theta(2)) \\
 &\tag{148}
 \end{aligned}$$

$$\begin{aligned}
 &= I(\mathbf{W}_{[K] \setminus \{\theta\}}; \mathbf{Z}'_{\mathcal{L}_1}(1), \mathbf{Z}'_{\mathcal{L}_2}(2), \mathbf{R}'_{\mathcal{H}_1}(1), \\
 &\quad \mathbf{R}'_{\mathcal{H}_2}(2) | \delta_1(1), \dots, \delta_B(2), \mathbf{W}_\theta(1), \mathbf{W}_\theta(2))
 \end{aligned} \tag{149}$$

$$= 0, \tag{150}$$

where (146) follows from the data processing inequality and the fact that the eavesdropper can also apply the same algorithm applied by the user to know the location of the Byzantine servers and their transmitted noise symbols, $(a) = 0$ since all messages are secure against any B servers, (149) is due to (127) and the independence of the noise variables, and (150) is due to independence between the masking variables and the messages and due to Lemma 6. ■

Remark 17: To see why it is enough to consider $(\mathbf{M}'_1 \mathbf{x})_{[N+1, 2N-2U]}$ for the answers: Considering that the servers share the quantum state and encoding operation done by them, the answers the user gets will be elements of an orthonormal basis for the whole quantum system of all servers.

This orthonormality then translates to perfect distinguishability in the measurements, which means that the measurement results completely describe the quantum answers the user has received. Thus, $(\mathbf{M}'_1 \mathbf{x})_{[N+1, 2N-2U]}$, which is the classical measurement result, is actually a description of the quantum answer the user has received.

Lemma 8: For the first regime in Theorem 3, the scheme is secure against the dynamic eavesdropper.

Proof: First, recall that, in the first regime, the number of eavesdropped links is $E \leq 2M + 2H + 2B - N$. In addition, for the vector $\mathbf{x}$, we have at least $2H + 2M + 2B$ independent random variables, i.e., $\mathbf{Z}'_{[H+M]}$, and $\mathbf{R}'_{[B]}$. In addition, we have

$$I(\mathbf{Z}'_{\mathcal{E}}(1); \hat{\mathbf{Z}}_B(1)) = 0, \tag{151}$$

where $|\mathcal{E}| \leq E$ and $H + M > B$.

Furthermore, $\mathbf{M}'_1 \mathbf{x}$ hides all the message symbols with $2M + 2H + 2B - N$ independent random variables from the interference symbols. Thus,

$$\begin{aligned}
 &I(\mathbf{W}_{[K]}; \theta; (\mathbf{M}'_1 \mathbf{x})_{\mathcal{E}_1}, \mathbf{Q}_{\mathcal{E}_2}) \\
 &= \underbrace{I(\theta; \mathbf{Q}_{\mathcal{E}_2})}_{=0 \text{ (by Lemma 4)}} + I(\theta; (\mathbf{M}'_1 \mathbf{x})_{\mathcal{E}_1} | \mathbf{Q}_{\mathcal{E}_2}) \\
 &\quad + \underbrace{I(\mathbf{W}_{[K]}; \mathbf{Q}_{\mathcal{E}_2} | \theta)}_{=0 \text{ (by assumption)}} + I(\mathbf{W}_{[K]}; (\mathbf{M}'_1 \mathbf{x})_{\mathcal{E}_1} | \mathbf{Q}_{\mathcal{E}_2}, \theta)
 \end{aligned} \tag{152}$$

$$\begin{aligned}
 &= I(\theta; \{Z'_i, (\mathbf{V}\mathbf{y})_j, \Delta_k(1), \Delta_\ell(2)\}_{\substack{i \in \mathcal{I}, j \in \mathcal{J}, k \in \mathcal{K}, \ell \in \mathcal{L} \\ |\mathcal{I} \cup \mathcal{J} \cup \mathcal{K} \cup \mathcal{L}| = |\mathcal{E}_1|}} | \mathbf{Q}_{\mathcal{E}_2}) \\
 &\quad + I(\mathbf{W}_{[K]}; \{Z'_i, (\mathbf{V}\mathbf{y})_j, \Delta_k(1), \Delta_\ell(2)\}_{\substack{i \in \mathcal{I}, j \in \mathcal{J}, k \in \mathcal{K}, \ell \in \mathcal{L} \\ |\mathcal{I} \cup \mathcal{J} \cup \mathcal{K} \cup \mathcal{L}| = |\mathcal{E}_1|}} | \mathbf{Q}_{\mathcal{E}_2}, \theta)
 \end{aligned} \tag{153}$$

$$\begin{aligned}
 &= I(\theta; \{\Delta_k(1), \Delta_\ell(2)\}_{\substack{k \in \mathcal{K}, \ell \in \mathcal{L} \\ |\mathcal{K} \cup \mathcal{L}| \leq \min\{4B, |\mathcal{E}_1| \cup \mathcal{I} \cup \mathcal{J}\}}} | \mathbf{Q}_{\mathcal{E}_2}) \\
 &\quad + I(\theta; \{Z'_i, (\mathbf{V}\mathbf{y})_j\}_{\substack{i \in \mathcal{I}, j \in \mathcal{J} \\ |\mathcal{I} \cup \mathcal{J}| \leq |\mathcal{E}_1|}} | \mathbf{Q}_{\mathcal{E}_2}, \{\Delta_k(1), \Delta_\ell(2)\}_{\substack{k \in \mathcal{K}, \ell \in \mathcal{L} \\ |\mathcal{K} \cup \mathcal{L}| \leq \min\{4B, |\mathcal{E}_1| \cup \mathcal{I} \cup \mathcal{J}\}}} | \mathbf{Q}_{\mathcal{E}_2}, \theta) \\
 &\quad + I(\mathbf{W}_{[K]}; \{\Delta_k(1), \Delta_\ell(2)\}_{\substack{k \in \mathcal{K}, \ell \in \mathcal{L} \\ |\mathcal{K} \cup \mathcal{L}| \leq \min\{4B, |\mathcal{E}_1| \cup \mathcal{I} \cup \mathcal{J}\}}} | \mathbf{Q}_{\mathcal{E}_2}, \theta) \\
 &\quad + I(\mathbf{W}_{[K]}; \{Z'_i, (\mathbf{V}\mathbf{y})_j\}_{\substack{i \in \mathcal{I}, j \in \mathcal{J} \\ |\mathcal{I} \cup \mathcal{J}| \leq |\mathcal{E}_1|}} | \mathbf{Q}_{\mathcal{E}_2}, \theta,
 \end{aligned} \tag{154}$$

$$\begin{aligned}
 &\quad \{\Delta_k(1), \Delta_\ell(2)\}_{\substack{k \in \mathcal{K}, \ell \in \mathcal{L} \\ |\mathcal{K} \cup \mathcal{L}| \leq \min\{4B, |\mathcal{E}_1| \cup \mathcal{I} \cup \mathcal{J}\}}}) \\
 &\leq I(\theta, \mathbf{Q}_B | \mathbf{Q}_{\mathcal{E}_2}) + I(\mathbf{W}_{[K]}; \mathbf{S}_B | \mathbf{Q}_{\mathcal{E}_2}, \theta) = 0,
 \end{aligned} \tag{155}$$

where $\mathcal{I} \subset [N]$, $\mathcal{J} \subset [N + 1 : N - 4B - 2U]$, $\mathcal{K} \subset [N - 4B - 2U + 1 : N - 2B - 2U]$, and $\mathcal{L} \subset [N - 2B - 2U + 1 : N - 2U]$, $\mathbf{V} = \mathbf{V}_{2L+2H+2M-N}(\mathbf{b}_{[2L+2H+2M+2B-N]})$, and

$$\mathbf{y} = \begin{bmatrix} \tilde{\mathbf{W}}_\theta(1) \\ \tilde{\mathbf{W}}_\theta(2) \\ \tilde{\mathbf{I}}_2(1) \\ \tilde{\mathbf{I}}_2(2) \end{bmatrix}. \tag{156}$$

Here, (155) follows from the data processing inequality, independence between the queries, messages and masking variables, storage security against any B servers and (151). ■

Now, for the second regime, we can write the transmitted symbols as

$$\mathbf{h}_2 = \begin{bmatrix} \mathbf{I}_N & 0 & 0 \\ 0 & \tilde{\mathbf{V}}_5 & 0 \\ 0 & 0 & \mathbf{I}_{4B+2U} \end{bmatrix} \underbrace{\begin{bmatrix} \tilde{\mathbf{I}}_1(1) \\ \tilde{\mathbf{I}}_1(2) \\ \tilde{\mathbf{r}}_{[\delta]} \\ \tilde{\mathbf{W}}_\theta(1) \\ \tilde{\mathbf{W}}_\theta(2) \\ \tilde{\mathbf{I}}_2(1) \\ \tilde{\mathbf{I}}_2(2) \\ \tilde{\Delta}_B(1) \\ \tilde{\Delta}_B(2) \\ \Delta'_U(1) \\ \Delta'_U(2) \end{bmatrix}}_{\mathbf{x}}, \quad (157)$$

where $\tilde{\mathbf{V}}_5 = \mathbf{V}_{L_1+L_2+\delta+2H+2M+2B-N}(\mathbf{b}_{[2L+2H+2M+2B-N]})$,

$$\begin{bmatrix} \tilde{\mathbf{r}}_{[\delta]} \\ \tilde{\mathbf{W}}_\theta(1) \\ \tilde{\mathbf{I}}_1(1) \\ \tilde{\mathbf{I}}_2(1) \\ \tilde{\Delta}_B(1) \end{bmatrix} = \begin{bmatrix} \mathbf{r}_{[\delta]} \\ \mathbf{W}_\theta(1) \\ \hat{\mathbf{I}}_{[H+M+B]}(1) \\ \mathbf{0}_{2B} \end{bmatrix} + \text{CSA}_{N-U,L,N-U}^{-1} \Delta_B(1), \quad (158)$$

$$\begin{bmatrix} \tilde{\mathbf{W}}_\theta(2) \\ \tilde{\mathbf{I}}_1(2) \\ \tilde{\mathbf{I}}_2(2) \\ \tilde{\Delta}_B(2) \end{bmatrix} = \begin{bmatrix} \mathbf{W}_\theta(2) \\ \hat{\mathbf{I}}_{[H+M+B]}(2) \\ \mathbf{0}_{2B} \end{bmatrix} + \text{CSA}_{N-U,L,N-U}^{-1} \Delta_B(2), \quad (159)$$

and

$$\hat{\mathbf{I}}_{[H+M+B]}(i) = \begin{bmatrix} \mathbf{I}_{[H+M]}(i) + \mathbf{Z}'_{[H+M]}(i) \\ \mathbf{R}'_{[B]}(i) \end{bmatrix}. \quad (160)$$

It is clear that the same approach can be used to prove symmetric privacy and eavesdropper protection.

Similarly, for the third regime, the transmitted symbols can be written as

$$\mathbf{h}_3 = \begin{bmatrix} \mathbf{I}_N & 0 & 0 \\ 0 & \mathbf{V}_{L_1+L_2+E}(\mathbf{b}_{[L_1+L_2+E]}) & 0 \\ 0 & 0 & \mathbf{I}_{4B+2U} \end{bmatrix} \underbrace{\begin{bmatrix} \tilde{\mathbf{I}}_1(1) \\ \tilde{\mathbf{I}}_1(2) \\ \tilde{\mathbf{r}}_{[E]} \\ \tilde{\mathbf{W}}_\theta(1) \\ \tilde{\mathbf{W}}_\theta(2) \\ \tilde{\Delta}_B(1) \\ \tilde{\Delta}_B(2) \\ \Delta'_U(1) \\ \Delta'_U(2) \end{bmatrix}}_{\mathbf{x}}, \quad (161)$$

where

$$\begin{bmatrix} \tilde{\mathbf{r}}_{[E]} \\ \tilde{\mathbf{W}}_\theta(1) \\ \tilde{\mathbf{I}}_1(1) \\ \tilde{\mathbf{I}}_2(1) \\ \tilde{\Delta}_B(1) \end{bmatrix} = \begin{bmatrix} \mathbf{r}_{[E]} \\ \mathbf{W}_\theta(1) \\ \hat{\mathbf{I}}_{[H+T_1+B]}(1) \\ \mathbf{0}_{2B} \end{bmatrix} + \text{CSA}_{N-U,L,N-U}^{-1} \Delta_B(1), \quad (162)$$

$$\begin{bmatrix} \tilde{\mathbf{W}}_\theta(2) \\ \tilde{\mathbf{I}}_1(2) \\ \tilde{\mathbf{I}}_2(2) \\ \tilde{\Delta}_B(2) \end{bmatrix} = \begin{bmatrix} \mathbf{W}_\theta(2) \\ \hat{\mathbf{I}}_{[H+T_2+B]}(2) \\ \mathbf{0}_{2B} \end{bmatrix} + \text{CSA}_{N-U,L,N-U}^{-1} \Delta_B(2), \quad (163)$$

and

$$\hat{\mathbf{I}}_{[H+T_i+B]}(i) = \begin{bmatrix} \mathbf{I}_{[H+T_i]}(i) + \mathbf{Z}'_{[H+T_i]}(i) \\ \mathbf{R}'_{[B]}(i) \end{bmatrix}. \quad (164)$$

This is also analogous to the previous case with $\delta = E$ and $2H + 2M + 2B = N$. Thus, the proofs for symmetric privacy and eavesdropper privacy go along the same lines.

Finally, for the last regime, recall that we can write the answers collectively from the $N - U$ serves as follows

$$\begin{bmatrix} A_{i_1}^{[\theta]} \\ \vdots \\ A_{i_{N-U}}^{[\theta]} \end{bmatrix} = \text{CSA}_{N-U,L,N-U} \mathbf{x}, \quad (165)$$

where

$$\mathbf{x} = \begin{bmatrix} \mathbf{W}_\theta \\ \mathbf{I}_{[H+M]} + \mathbf{Z}'_{[H+M]} \\ \mathbf{R}'_{[B]} \\ \mathbf{0}_{2B} \end{bmatrix} + \text{CSA}_{N-U,L,N-U}^{-1} \Delta_B. \quad (166)$$

In this regime as well, we can see that, by using the same approach that is used in the first regime, security and privacy requirements are satisfied.

Remark 18: Using the result from [36], we know that the transfer matrix defined by

$$[\mathbf{0}_N \ \mathbf{I}_N] \mathbf{M}^{-1}, \quad (167)$$

where $\mathbf{M}$ is as defined in (40) or (75), is a valid N -sum box transfer matrix.

Lemma 9: Let $\mathbf{D}_1 = [\mathbf{0} \ \mathbf{I}_N] [\mathbf{G} \ \mathbf{H}]^{-1}$ be a feasible N -sum box transfer matrix, then for any $N \times N$ invertible matrices $\mathbf{V}_1$ and $\mathbf{V}_2$, $\mathbf{D}_2 = [\mathbf{0} \ \mathbf{I}_N] \begin{bmatrix} \mathbf{V}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 \end{bmatrix}^{-1} [\mathbf{G} \ \mathbf{H}]^{-1}$ is a feasible N -sum box transfer matrix.

Proof: Note that since both $\mathbf{V}_1$ and $\mathbf{V}_2$ are invertible, then $\begin{bmatrix} \mathbf{V}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 \end{bmatrix}$ is also invertible. Now, since $\mathbf{G}$ is an SSO matrix, we need to check that $\mathbf{G}\mathbf{V}_1$ is also an SSO matrix since

$$\mathbf{D}_2 = [\mathbf{0} \ \mathbf{I}_N] \begin{bmatrix} \mathbf{V}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 \end{bmatrix}^{-1} [\mathbf{G} \ \mathbf{H}]^{-1} \quad (168)$$

$$= [\mathbf{0} \ \mathbf{I}_N] \left([\mathbf{G} \ \mathbf{H}] \begin{bmatrix} \mathbf{V}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 \end{bmatrix} \right)^{-1} \quad (169)$$

$$= [\mathbf{0} \ \mathbf{I}_N] ([\mathbf{G}\mathbf{V}_1 \ \mathbf{H}\mathbf{V}_2])^{-1}. \quad (170)$$

Thus, $\mathbf{V}_1^\dagger \mathbf{G}^\dagger \mathbf{G} \mathbf{V}_1 = \mathbf{V}_1^\dagger (\mathbf{G}^\dagger \mathbf{J} \mathbf{G}) \mathbf{V}_1 = \mathbf{0}$ implying that $\mathbf{G}\mathbf{V}_1$ is a valid SSO matrix. Now, since $[\mathbf{G} \ \mathbf{H}]$ is invertible, thus $[\mathbf{G}\mathbf{V}_1 \ \mathbf{H}\mathbf{V}_2] = [\mathbf{G} \ \mathbf{H}] \begin{bmatrix} \mathbf{V}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 \end{bmatrix}$ is also invertible. ■

REFERENCES

- [1] K. Banawan and S. Ulukus, "The capacity of private information retrieval from Byzantine and colluding databases," *IEEE Trans. Inf. Theory*, vol. 65, no. 2, pp. 1206–1219, Feb. 2019.
- [2] S. Song and M. Hayashi, "Capacity of quantum private information retrieval with multiple servers," *IEEE Trans. Inf. Theory*, vol. 67, no. 1, pp. 452–463, Jan. 2021.
- [3] S. Song and M. Hayashi, "Capacity of quantum private information retrieval with colluding servers," *IEEE Trans. Inf. Theory*, vol. 67, no. 8, pp. 5491–5508, Aug. 2021.
- [4] M. Allaix, Y. Lu, Y. Yao, T. Pllaha, C. Hollanti, and S. A. Jafar, " N -sum box: An abstraction for linear computation over many-to-one quantum networks," *IEEE Trans. Info. Theory*, vol. 71, no. 2, pp. 1121–1139, Dec. 2025.
- [5] Q. Wang, H. Sun, and M. Skoglund, "The capacity of private information retrieval with eavesdroppers," *IEEE Trans. Inf. Theory*, vol. 65, no. 5, pp. 3198–3214, May 2019.

- [6] K. Banawan and S. Ulukus, "Private information retrieval through wiretap channel II: Privacy meets security," *IEEE Trans. Inf. Theory*, vol. 66, no. 7, pp. 4129–4149, Jul. 2020.
- [7] A. Aytakin, M. Nomeir, S. Vithana, and S. Ulukus, "Quantum X-secure E-eavesdropped T-colluding symmetric private information retrieval," *IEEE Trans. Inf. Theory*, vol. 71, no. 5, pp. 3974–3988, May 2025.
- [8] B. Chor, E. Kushilevitz, O. Goldreich, and M. Sudan, "Private information retrieval," *J. ACM*, vol. 45, no. 6, pp. 965–981, Nov. 1998.
- [9] H. Sun and S. A. Jafar, "The capacity of private information retrieval," *IEEE Trans. Inf. Theory*, vol. 63, no. 7, pp. 4075–4088, Jul. 2017.
- [10] H. Sun and S. A. Jafar, "The capacity of symmetric private information retrieval," *IEEE Trans. Inf. Theory*, vol. 65, no. 1, pp. 322–329, Jan. 2019.
- [11] H. Sun and S. A. Jafar, "The capacity of robust private information retrieval with colluding databases," *IEEE Trans. Inf. Theory*, vol. 64, no. 4, pp. 2361–2370, Apr. 2018.
- [12] Q. Wang and M. Skoglund, "Symmetric private information retrieval from MDS coded distributed storage with non-colluding and colluding servers," *IEEE Trans. Inf. Theory*, vol. 65, no. 8, pp. 5160–5175, Aug. 2019.
- [13] X. Yao, N. Liu, and W. Kang, "The capacity of private information retrieval under arbitrary collusion patterns for replicated databases," *IEEE Trans. Inf. Theory*, vol. 67, no. 10, pp. 6841–6855, Oct. 2021.
- [14] Q. Wang and M. Skoglund, "On PIR and symmetric PIR from colluding databases with adversaries and eavesdroppers," *IEEE Trans. Inf. Theory*, vol. 65, no. 5, pp. 3183–3197, May 2019.
- [15] R. Tajeddine, O. W. Gnilke, D. Karpuk, R. Freij-Hollanti, and C. Hollanti, "Private information retrieval from coded storage systems with colluding, Byzantine, and unresponsive servers," *IEEE Trans. Inf. Theory*, vol. 65, no. 6, pp. 3898–3906, Jun. 2019.
- [16] H. Yang, W. Shin, and J. Lee, "Private information retrieval for secure distributed storage systems," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 12, pp. 2953–2964, Dec. 2018.
- [17] Z. Jia, H. Sun, and S. A. Jafar, "Cross subspace alignment and the asymptotic capacity of X-secure T-private information retrieval," *IEEE Trans. Inf. Theory*, vol. 65, no. 9, pp. 5783–5798, Sep. 2019.
- [18] Z. Jia and S. A. Jafar, "X-secure T-private information retrieval from MDS coded storage with Byzantine and unresponsive servers," *IEEE Trans. Inf. Theory*, vol. 66, no. 12, pp. 7427–7438, Dec. 2020.
- [19] K. Banawan and S. Ulukus, "Multi-message private information retrieval: Capacity results and near-optimal schemes," *IEEE Trans. Inf. Theory*, vol. 64, no. 10, pp. 6842–6862, Oct. 2018.
- [20] K. Banawan, B. Arasli, Y.-P. Wei, and S. Ulukus, "The capacity of private information retrieval from heterogeneous uncoded caching databases," *IEEE Trans. Inf. Theory*, vol. 66, no. 6, pp. 3407–3416, Jun. 2020.
- [21] K. Banawan and S. Ulukus, "The capacity of private information retrieval from coded databases," *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 1945–1956, Mar. 2018.
- [22] C. Tian, H. Sun, and J. Chen, "Capacity-achieving private information retrieval codes with optimal message size and upload cost," *IEEE Trans. Inf. Theory*, vol. 65, no. 11, pp. 7613–7627, Nov. 2019.
- [23] L. Holzbaur, R. Freij-Hollanti, and C. Hollanti, "On the capacity of private information retrieval from coded, colluding, and adversarial servers," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Aug. 2019, pp. 1–5.
- [24] N. Raviv, I. Tamo, and E. Yaakobi, "Private information retrieval in graph-based replication systems," *IEEE Trans. Inf. Theory*, vol. 66, no. 6, pp. 3590–3602, Jun. 2020.
- [25] Y.-P. Wei, K. Banawan, and S. Ulukus, "Fundamental limits of cache-aided private information retrieval with unknown and uncoded prefetching," *IEEE Trans. Inf. Theory*, vol. 65, no. 5, pp. 3215–3232, May 2019.
- [26] M. Nomeir, S. Vithana, and S. Ulukus, "Asymmetric X-secure T-private information retrieval: More databases is not always better," in *Proc. 58th Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2024, pp. 1–6.
- [27] M. A. Attia, D. Kumar, and R. Tandon, "The capacity of private information retrieval from uncoded storage constrained databases," *IEEE Trans. Inf. Theory*, vol. 66, no. 11, pp. 6617–6634, Nov. 2020.
- [28] S. Ulukus, S. Avestimehr, M. Gastpar, S. A. Jafar, R. Tandon, and C. Tian, "Private retrieval, computing, and learning: Recent progress and future challenges," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 3, pp. 729–748, Mar. 2022.
- [29] S. Vithana, K. Banawan, and S. Ulukus, "Semantic private information retrieval," *IEEE Trans. Inf. Theory*, vol. 68, no. 4, pp. 2635–2652, Apr. 2022.
- [30] A. Heidarzadeh, S. Kadhe, S. El Rouayheb, and A. Sprintson, "Single-server multi-message individually-private information retrieval with side information," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2019, pp. 1042–1046.
- [31] Z. Chen, Z. Wang, and S. A. Jafar, "The capacity of T-private information retrieval with private side information," *IEEE Trans. Inf. Theory*, vol. 66, no. 8, pp. 4761–4773, Aug. 2020.
- [32] R. Tajeddine, O. W. Gnilke, and S. El Rouayheb, "Private information retrieval from MDS coded data in distributed storage systems," *IEEE Trans. Inf. Theory*, vol. 64, no. 11, pp. 7081–7093, Nov. 2018.
- [33] O. Makkonen, D. A. Karpuk, and C. Hollanti, "Algebraic geometry codes for cross-subspace alignment in private information retrieval," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2024, pp. 2874–2879.
- [34] O. Makkonen, D. Karpuk, and C. Hollanti, "Secret sharing for secure and private information retrieval: A construction using algebraic geometry codes," 2024, *arXiv:2408.00542*.
- [35] A. Aytakin, M. Nomeir, S. Vithana, and S. Ulukus, "Quantum symmetric private information retrieval with secure storage and eavesdroppers," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2023, pp. 1057–1062.
- [36] Y. Lu and S. A. Jafar, "Quantum X-secure T-private information retrieval from MDS coded storage with unresponsive and Byzantine servers," 2023, *arXiv:2311.07829*.
- [37] M. Nomeir, A. Aytakin, and S. Ulukus, "Quantum X-secure B-Byzantine T-colluding private information retrieval," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Nov. 2024, pp. 705–710.
- [38] Y. Lu, Y. Yao, and S. A. Jafar, "On the capacity of secure K-user product computation over a quantum MAC," *IEEE Commun. Lett.*, vol. 27, no. 10, pp. 2598–2602, Oct. 2023.
- [39] Y. Yao and S. A. Jafar, "The capacity of classical summation over a quantum MAC with arbitrarily distributed inputs and entanglements," *IEEE Trans. Inf. Theory*, vol. 70, no. 9, pp. 6350–6370, Sep. 2024.
- [40] Y. Yao and S. A. Jafar, "On the capacity of vector linear computation over a noiseless quantum multiple access channel with entangled transmitters," 2024, *arXiv:2407.01498*.
- [41] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 1999.
- [42] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge, U.K.: Cambridge Univ. Press, 2010.
- [43] M. Hayashi, *Group Representation for Quantum Theory*. Cham, Switzerland: Springer, 2017.
- [44] Y. Lu and S. A. Jafar, "Quantum cross subspace alignment codes via the N-sum box abstraction," in *Proc. 57th Asilomar Conf. Signals, Syst., Comput.*, Oct. 2023, pp. 670–674.
- [45] D. Gottesman. (2024). *Surviving as a Quantum Computer in a Classical World*. [Online]. Available: <https://www.cs.umd.edu/class/spring2024/cmssc858G/QECCbook-2024-ch1-15.pdf>
- [46] A. Aytakin, M. Nomeir, and S. Ulukus. (2024). *Quantum X-Secure B-Byzantine T-Colluding Private Information Retrieval*. [Online]. Available: <http://www.ece.umd.edu/~ulukus/papers/journal/QXBTPIR.pdf>

Mohamed Nomeir (Graduate Student Member, IEEE) received the B.S. degree in electrical engineering from Alexandria University, Egypt, and the M.Sc. degree from the Electronics and Communications Department, The American University in Cairo, Egypt. He is currently pursuing the Ph.D. degree with the ECE Department, UMD. His current research interests include classical and quantum information theoretic privacy and security. He received the Medhat Haroun Award for Scholarly Excellence as the best thesis of the academic year for his master's thesis.

Alptug Aytakin (Graduate Student Member, IEEE) received the bachelor's degree in electrical engineering and physics from Bilkent University. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, USA. His research interests include classical and quantum information theory.

Sennur Ulukus (Fellow, IEEE) received the B.S. and M.S. degrees in electrical and electronics engineering from Bilkent University and the Ph.D. degree in electrical and computer engineering from the Wireless Information Network Laboratory (WINLAB), Rutgers University.

She is currently a Distinguished University Professor and the Anthony Ephremides Professor in information sciences and systems with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, USA, where she also holds a joint appointment with the Institute for Systems Research (ISR). She is also a Distinguished Scholar-Teacher with the University of Maryland. Prior to joining UMD, she was a Senior Technical Staff Member with AT&T Laboratories-Research. Her research interests include information theory, wireless communications, machine learning, signal processing and networks, with recent focus on age of information, private information retrieval, blockchain consensus, quantum information, machine learning for wireless, federated learning, and large language models for communications and networks.

Dr. Ulukus received the 2003 IEEE Marconi Prize Paper Award in Wireless Communications, the 2005 NSF CAREER Award, the 2011 ISR Outstanding Systems Engineering Faculty Award, the 2012 ECE George Corcoran Outstanding Teaching Award, the 2019 IEEE Communications Society Best Tutorial Paper Award, the 2020 IEEE Communications Society Women in Communications Engineering (WICE) Outstanding Achievement Award, the 2020 IEEE Communications Society Technical Committee on Green Communications and Computing (TCGCC) Distinguished Technical Achievement Recognition Award, and the 2024 IEEE Communications Society Communication Theory Technical Committee (CTTC) Technical Achievement Award. She was the TPC Chair of 2021 IEEE Globecom and the TPC Co-Chair of 2011 IEEE CTW, 2014 IEEE PIMRC, 2016 IEEE Globecom, 2017 IEEE ISIT, 2019 IEEE ITW, 2023 IEEE MILCOM, 2024 IEEE DySPAN, and 2024 IEEE Globecom. She has been a Senior Editor of IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING since 2020. She was an Area Editor of IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING from 2016 to 2020 and IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2019 to 2023. She was an Editor of IEEE TRANSACTIONS ON COMMUNICATIONS from 2003 to 2007 and IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS Series on Green Communications and Networking from 2015 to 2016. She was an Associate Editor of IEEE TRANSACTIONS ON INFORMATION THEORY from 2007 to 2010. She was a Guest Editor of IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS in 2008, 2015, 2021, and 2022; IEEE JOURNAL ON SELECTED AREAS IN INFORMATION THEORY in 2021, 2022, and 2023; IEEE TRANSACTIONS ON INFORMATION THEORY in 2011; *Journal of Communications and Networks* in 2012; and IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING in 2021. She was a Distinguished Lecturer of the IEEE Information Theory Society from 2018 to 2019.