


Quantum Byzantine Agreement Against Full-Information Adversary

Longcheng Li ✉ 


State Key Lab of Processors, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China

School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing, China

Xiaoming Sun ✉ 

State Key Lab of Processors, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China

School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing, China

Jiadong Zhu¹ ✉ 

State Key Lab of Processors, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China

Abstract

We exhibit that, when given a classical Byzantine agreement protocol designed in the private-channel model, it is feasible to construct a quantum agreement protocol that can effectively handle a full-information adversary. Notably, both protocols have equivalent levels of resilience, round complexity, and communication complexity. In the classical private-channel scenario, participating players are limited to exchanging classical bits, with the adversary lacking knowledge of the exchanged messages. In contrast, in the quantum full-information setting, participating players can exchange qubits, while the adversary possesses comprehensive and accurate visibility into the system's state and messages. By showcasing the reduction from quantum to classical frameworks, this paper demonstrates the strength and flexibility of quantum protocols in addressing security challenges posed by adversaries with increased visibility. It underscores the potential of leveraging quantum principles to improve security measures without compromising on efficiency or resilience.

By applying our reduction, we demonstrate quantum advantages in the round complexity of asynchronous Byzantine agreement protocols in the full-information model. It is well known that in the full-information model, any classical protocol requires $\Omega(n)$ rounds to solve Byzantine agreement with probability one even against Fail-stop adversary when resilience $t = \Theta(n)$ [2]. We show that quantum protocols can achieve $O(1)$ rounds (i) with resilience $t < n/2$ against a Fail-stop adversary, and (ii) with resilience $t < n/(3 + \epsilon)$ against a Byzantine adversary for any constant $\epsilon > 0$, therefore surpassing the classical lower bound.

2012 ACM Subject Classification Theory of computation → Distributed algorithms; Theory of computation → Quantum computation theory

Keywords and phrases Byzantine agreement, Quantum computation, Full-information model

Digital Object Identifier 10.4230/LIPIcs.DISC.2024.32

Related Version *Extended Version*: <https://arxiv.org/abs/2409.01707>

Funding This work was supported in part by the National Natural Science Foundation of China Grants No. 62325210, and the Strategic Priority Research Program of Chinese Academy of Sciences Grant No. XDB28000000.

¹ Corresponding author



1 Introduction

Byzantine agreement (BA) [32], also referred to as Byzantine fault-tolerant distributed consensus, is a crucial topic in secure distributed computing. In simple terms, in a BA protocol, a group of n players who do not trust each other and possess private input bits, come to a consensus on a shared output bit, even if a subset of size t of the players are corrupted by a malicious adversary, who can force the corrupt parties to deviate from their prescribed programs during the protocol execution. The Byzantine agreement problem has been extensively researched over the past four decades, leading to numerous findings on the feasibility and potential of BA protocols in various settings [19, 11, 3].

In this paper, we focus on BA that succeeds with probability one in the full-information model, where the adversary knows the knowledge of all local variables, including quantum states if applicable. It is well known that in this model, when up to t players may be corrupted, no classical deterministic protocol can solve synchronous BA in less than $t + 1$ rounds even in the presence of a Fail-stop adversary [32]. It is further proved by [5] that any classical randomized protocol requires at least expected $\tilde{\Omega}(\sqrt{n})$ rounds. Given these constraints, it is natural to ask the following question:

Can quantum communication accelerate BA in the full-information model?

The seminal work of [7] provides a confirming answer to the above question by constructing a constant round synchronous quantum BA protocol against the Byzantine adversary, surpassing the established round complexity lower bound in [5]. The protocol builds upon an expected constant round classical BA protocol introduced in [19], which is not resilient against a full-information adversary and requires a private channel. In their work, [7] proposes a quantum modification to the original classical protocol to make it robust against a full-information adversary. They achieve this by introducing a novel approach of deferring coin flips, substituting them with quantum superpositions until after the adversary has chosen his actions in a certain round. Notably, the modification does not change the structure of the original classical protocol and therefore preserves its constant round complexity. [7] also extends the synchronous quantum protocol to the asynchronous case, but with suboptimal resilience $t < n/4$.

[7] demonstrates an elegant method of reducing quantum full-information protocols to classical private-channel protocols while maintaining key attributes such as resilience, round complexity, and communication complexity. This approach offers a valuable means of evaluating the quantum advantage in the full-information model by comparing classical full-information and classical private-channel models. By highlighting the notable distinctions between these two classical models, they underscore the substantial quantum advantage inherent in the full-information domain.

In light of these findings, [7] raises the question of whether their reduction strategy could be applied to other settings, such as low round complexity asynchronous BA protocols with resilience $n/4 \leq t < n/3$, to further investigate potential quantum advantages. Unfortunately, limited progress has been made on this issue since its introduction. This paper seeks to tackle this challenge from a comprehensive viewpoint. Instead of narrowly focusing on the reduction of quantum protocols to classical protocols in a specific setting (e.g., the asynchronous protocol with resilience $n/4 \leq t < n/3$, which is better than that in [7]), our objective is to address the following question:

*Is it possible to convert **any** classical private-channel BA protocol to a quantum full-information BA protocol while preserving the same characteristics such as resilience, round complexity, and communication complexity?*

■ **Table 1** Round complexity of Byzantine agreement in the full-information model.

Model	Adversary	Resilience	Classical		Quantum ^{a)}	
			Upper bound	Lower bound	Upper bound	\mathcal{P}_C
Sync.	Fail-stop	$t = \Theta(n)$	$\tilde{O}(\sqrt{n})$ [5]	$\tilde{\Omega}(\sqrt{n})$ [5]	$O(1)$ [7, 24]	[16]
	Byzantine	$t < n/3$	$O(n)$ [32]		$O(1)$ [7]	[19]
Async.	Fail-stop	$t = \Theta(n)$	$O(n)$ [2]		$O(1)$ (Our work)	[3]
	Byzantine	$t < n/4$	$\tilde{O}(n^4)$ [26]	$\Omega(n)$ [2]	$O(1)$ [7]	[19]
	Byzantine	$t < \frac{n}{3+\epsilon}$ ^{b)}	$\tilde{O}(n^4/\epsilon^8)$ [26]		$O(1/\epsilon)$ (Our work)	[4]
	Byzantine	$t < n/3$	$\tilde{O}(n^{12})$ [26]		$O(n)$ (Our work)	[4]

a) Every quantum protocol presented in the table is built upon some classical private-channel protocol \mathcal{P}_C . The last two columns of the table show the classical private-channel protocols alongside their quantum full-information equivalents for comparison and reference purposes.

b) Notice that when ϵ is a constant, the quantum upper bound is $O(1/\epsilon) = O(1)$.

1.1 Our Contribution

As our main result, we answer the above question in the affirmative by demonstrating a general reduction from a quantum full-information BA protocol to a classical private-channel BA protocol:

► **Theorem 1.** *Given a classical synchronous (resp. asynchronous) non-erasing BA protocol designed to counter a private-channel Fail-stop (resp. Byzantine) adversary, we can construct a quantum synchronous (resp. asynchronous) BA protocol capable of handling a full-information Fail-stop (resp. Byzantine) adversary while maintaining the same levels of resilience, round complexity, and communication complexity.*

It is crucial to emphasize that the theorem we present is applicable under the condition that the classical protocol forming the foundation of our quantum protocol is *non-erasing*, which means its security does not rely on the erasure of intermediate states. To the best of our knowledge, this criterion is met by all existing classical protocols within the scope of information-theoretic BA with probability one. For a more detailed definition of this concept, please refer to Definition 3 in Section 4 where we will provide a formal explanation. Furthermore, throughout our paper, we consistently assume that the adversary is computationally unlimited and adaptive¹, allowing it to modify its strategy based on the information acquired during the execution of the protocol.

By applying our reduction, we obtain several new quantum advantages related to round complexity in the full-information setting. As summarized in Table 1, our main result enables us to quantize existing classical private-channel protocols into some quantum full-information protocols of which the round complexity surpasses the classical lower bound in the same setting. In particular, we obtain two new quantum speedups in the asynchronous model:

■ **Fail-stop model:** Section 14.3 of [3] presents a constant-round classical BA protocol with optimal resilience $t < n/2$ against the Fail-stop adversary in the private-channel setting. By applying our reduction, we obtain a constant-round quantum full-information BA protocol with $t < n/2$, while any classical full-information protocol requires $\Omega(n)$ rounds [2].

¹ Similar reductions can also be made from the quantum non-adaptive full-information model to the classical non-adaptive private-channel model.

- **Byzantine model:** For any $\epsilon > 0$, [4] presents an $O(1/\epsilon)$ -round classical BA protocol with resilience $t < n/(3 + \epsilon)$ against the private-channel Byzantine adversary. By applying our reduction, we obtain an $O(1/\epsilon)$ -round quantum full-information BA protocol with resilience $t < n/(3 + \epsilon)$. When ϵ is a constant independent of n , the quantum BA achieves constant rounds, while any classical full-information protocol requires $\Omega(n)$ rounds [2]. When $\epsilon \leq 1/n$, $\lceil n/(3 + \epsilon) \rceil = \lceil n/3 \rceil$, which indicates that $t < n/(3 + \epsilon)$ is equivalent to $t < n/3$. By substituting $\epsilon = 1/n$ into $O(1/\epsilon)$, we find that our quantum BA requires $O(1/\epsilon) = O(n)$ rounds. In comparison, the best known classical protocol [26] in the same setting requires $\tilde{O}(n^{12})$ rounds.

1.2 Technical Overview

We briefly explain the key ideas behind Theorem 1, especially how to quantize a classical protocol into a quantum one and how to simulate a quantum full-information adversary in the classical setting. The key idea is utilizing quantum superpositions to turn exposed randomness into hidden randomness.

A simple motivating example. Before introducing the complicated quantum full-information BA protocol against the Byzantine adversary, [7] first presents a simple quantum full-information BA protocol against the Fail-stop adversary, who can corrupt players by halting it and choosing a subset of their messages to be delivered. This simple protocol follows a common framework of reducing a BA protocol to a common-coin protocol, where all uncorrupted players need to output a common random coin with constant success probability. We will use the common-coin protocol, as demonstrated in the BA protocol against the Fail-stop adversary in [7], as a motivating example to explain the key idea of our paper. The common-coin protocol works in the quantum full-information setting and draws inspiration from a common-coin protocol in the classical private-channel setting [16]. In the following discussion, we will start by offering a brief overview of the classical private-channel protocol in [16] and explaining its limitations when confronted with a full-information adversary. We then explain how [7] effectively resolve this issue by leveraging quantum principles.

The classical private-channel protocol in [16] works as follows: (i) Each player i picks a random coin $c_i \in \{0, 1\}$ and a random leader value $l_i \in [n^3]$ and then multicasts (c_i, l_i) ; (ii) Each player i outputs the coin c_j such that l_j is the largest leader value i receives. A private-channel Fail-stop adversary learns nothing about the values of $\{c_i\}$ and $\{l_i\}$, so the best it can do is to randomly stop t players. Since there are at least $n - t > n/2$ uncorrupted players, the largest leader falls among uncorrupted players with probability $1/2$, and the probability of collision of leader values is negligible. Switching to full-information adversary, $\{c_i\}$ and $\{l_i\}$ become known to the adversary. Then the adversary can corrupt the leader and let only a subset of players receive the leader's message so that it can break the common-coin protocol. However, [7] shows that the problem can be fixed if we allow quantumness. Instead of choosing random c_i and l_i , we let player i *purify randomness*, i.e., preparing two n -qudit superposition states

$$|c_i\rangle := \frac{1}{\sqrt{2}} (|00 \cdots 0\rangle + |11 \cdots 1\rangle) \quad \text{and} \quad |l_i\rangle := \frac{1}{\sqrt{n^3}} \sum_{l=1}^{n^3} |l, l, \dots, l\rangle,$$

and then distribute the n qudits of $|c_i\rangle$ and $|l_i\rangle$ among the players. In the next round, the players measure the qudits they receive and obtain the classical random coins and leader values. Although the full-information adversary can see the pure state of the system, quantum mechanics prevents it from knowing the random values before measurement. Thus this simple purified quantum protocol works against the full-information adversary.

Generalized reduction in the synchronous model. Inspired by the above example, we give a general reduction from quantum full-information BA protocols to classical private-channel BA protocols. For any classical BA protocol \mathcal{P}_C , the local computation of each player at round k involves (i) preparing some randomness r_k , and (ii) computing a function f to determine the decided value and messages to be sent. We construct a quantum protocol \mathcal{P}_Q by modifying \mathcal{P}_C 's local computation to (i) preparing a quantum state $\sum_r \sqrt{\Pr[r_k = r]} |r\rangle$, (ii) applying a unitary U_f to compute f reversibly i.e., $U_f |v\rangle |0\rangle := |v\rangle |f(v)\rangle$ and send quantum messages.

We assume that the output of f contains a variable $d_k \in \{0, 1, \perp\}$ indicating the decided value at round k (\perp if not decided yet). The player in \mathcal{P}_Q will measure the corresponding quantum register of d_k and decide if $d_k \neq \perp$. In addition, to prevent a communication blowup, we also assume the output of f includes the message pattern $b_k \in \{0, 1\}^n$ where the j -th bit $b_k[j]$ indicates whether to send message to player j . \mathcal{P}_Q will measure the register of b_k and send messages only to players with $b_k[j] = 1$.

Security analysis. To prove that \mathcal{P}_Q is secure against a quantum full-information adversary, we follow the argument that given any quantum full-information adversary \mathcal{A}_Q attacking \mathcal{P}_Q , we can construct a classical adversary \mathcal{A}_C in the private-channel model that perfectly simulates $(\mathcal{P}_Q, \mathcal{A}_Q)$ when interacting with \mathcal{P}_C . However, one may question why this simulation is possible since \mathcal{A}_Q is apparently more powerful than \mathcal{A}_C in two aspects:

1. \mathcal{A}_Q is full-information while \mathcal{A}_C is private-channel.
2. \mathcal{A}_Q is quantum while \mathcal{A}_C is classical.

For the first problem, observe that the randomness of \mathcal{P}_Q comes solely from players' measurement results of $\{b_k\}$ and $\{d_k\}$, of which the corresponding classical variables in \mathcal{P}_C are also available to \mathcal{A}_C .² The pure state view of \mathcal{P}_Q is fully determined by $\{b_k\}$ and $\{d_k\}$, so actually \mathcal{A}_Q knows no more than \mathcal{A}_C about the state of the system.

For the second problem, we first consider the Fail-stop adversary case to demonstrate why it is not a concern. The ability of a Fail-stop adversary \mathcal{A}_Q is to halt players and choose a subset of their messages to be delivered, which is essentially classical. Thus \mathcal{A}_C can easily simulate those actions.

The Byzantine case is trickier because a Byzantine adversary \mathcal{A}_Q can apply quantum operations on the registers of corrupted players. In this case, we let \mathcal{A}_C *classically simulate*³ a quantum state on the registers of corrupted players in order to keep track of \mathcal{A}_Q 's actions. Moreover, when corrupted players (controlled by \mathcal{A}_C) send messages to uncorrupted players, they cannot simply transmit quantum messages in the manner \mathcal{A}_Q does because players in \mathcal{P}_C are not equipped to receive quantum information. To circumvent this challenge, we let corrupted players first measure the messages and then send the measurement outcomes, which are classical, to the uncorrupted players. Intuitively, measuring those messages will not affect the simulation because uncorrupted players always keep a copy of messages they receive. After a quantum message is sent to an uncorrupted player, corrupted players are unable to reobtain it, resulting in the message being traced out from the corrupted players' system, which is equivalent to being measured. There is still one caveat in the simulation of \mathcal{A}_Q by \mathcal{A}_C : because \mathcal{A}_Q is adaptive, it can corrupt new players during the protocol and

² Private-channel \mathcal{A}_C knows message patterns by definition. We can also assume \mathcal{A}_C knows the decided values of players because if an uncorrupted player decides in a BA protocol, all other uncorrupted players will eventually decide the same value.

³ We assume the adversary is computationally unbounded.

reobtain the quantum messages sent to them previously, while \mathcal{A}_C will only obtain collapsed classical messages when corrupting new players. To fix this, we let \mathcal{A}_C maintain a copy T of the communication transcript between uncorrupted players and corrupted players. By following this approach, when \mathcal{A}_C corrupts some new players, it can replicate the necessary quantum states as per the content stored in T . In this way, \mathcal{A}_C can perfectly simulate \mathcal{A}_Q in the classical setting.

Round and communication complexity. Our construction of \mathcal{A}_C actually yields a stronger result: the probability distribution of executions in $(\mathcal{P}_Q, \mathcal{A}_Q)$ is identical to that of executions in $(\mathcal{P}_C, \mathcal{A}_C)$. This leads to the conclusion presented in Theorem 1.

Extending to the asynchronous model. Our results in the synchronous model can be extended to the asynchronous model without extra effort. The primary distinction lies in the measurement metric used; while synchronous protocols are evaluated in *rounds*, asynchronous protocols are evaluated in terms of *steps*. In one step, only one uncorrupted player receives a message, then performs local computation and possibly sends out messages. It is still feasible to purify the randomness, perform reversible computation in each step, and develop a quantum full-information protocol.

Although our results are inspired by the Fail-stop protocol in [7], our techniques are new compared with [7], especially in the Byzantine model. In the Byzantine model, [7] involves an intricate procedure of modifying the original classical protocol by replacing its classical verifiable secret sharing (VSS) component with a quantum VSS. In contrast, our approach focuses on demonstrating the efficacy of extracting purified classical randomness, a feature that is applicable to any classical protocol exhibiting a non-erasing property. Therefore, we expect our technique to have a broader range of applications.

2 Related Work

We address the construction of a quantum full-information protocol from a classical private-channel protocol. In this section, we discuss existing results in closely related contexts and provide a brief overview of their techniques.

BA protocols with private channels. The private-channel model is frequently studied in BA problems. In this model, the adversary is unable to access the contents of the messages exchanged between the participating players. A seminal work [19] presents a synchronous BA protocol that can withstand up to $t < n/3$ failures and operates within an expected constant number of rounds. Additional randomized protocols [35, 11] addressing scenarios where $n/3 \leq t < n/2$ are known, which require extra assumptions like a public-key infrastructure and a trusted dealer. Due to their dependency on these supplementary assumptions, these protocols cannot be adapted to the information-theoretic setting. In the information-theoretic setting, [3] presents an asynchronous BA protocol that can withstand up to $t < n/2$ failures while maintaining a constant running time, particularly effective against the Fail-stop adversary. For the Byzantine adversary, [1] introduces a concept called shunning verifiable secret sharing and gives an asynchronous BA protocol with optimal resilience $t < n/3$ and $O(n^2)$ running time, which is later improved to $O(n)$ by [4].

The full-information model. The full-information model, as introduced by [8], serves as a framework for investigating collective coin-flipping within a network of n players with t failures. This model has spurred a series of research efforts aimed at enhancing fault tolerance

and reducing round complexity in protocols such as those proposed by [34] and [18]. [23] considers the problem of multiparty computation in the full-information model. [27] gives the first asynchronous leader election protocol in the full information model with constant success probability against a constant fraction of corrupted players. Asynchronous BA in the full-information model used to require exponential time to be solved with linear resilience [6, 10], which is recently improved to polynomial time by a sequence of works [29, 25, 26].

Quantum Byzantine protocols. Besides the work of [7], many works have applied quantum principles to Byzantine fault tolerance problems, which has led to significant advancements in the field. A key contribution is made by [20], who introduces quantum elements to Byzantine problems by addressing a weaker version called Detectable Byzantine Agreement (DBA). Their protocol involves three parties and is based on the Aharonov state. Building upon this work, [22] proposes a 3-party DBA protocol utilizing four-particle entangled qubits. Further research by [21] shows that the DBA protocol can reach any tolerance found. Other variants of the problem setting [15, 30, 33] are considered to ensure feasibility of the problem against strong Byzantine adversaries. It is also worth mentioning that a recent work [24] improves the communication complexity of the synchronous Fail-stop protocol of [7] from $O(n^2)$ to $O(n^{1+\epsilon})$ for any constant $\epsilon > 0$ while maintaining constant running time.

3 Preliminaries

3.1 Quantum Computation

In this section, we will briefly discuss quantum computation. For a more in-depth explanation, readers are encouraged to refer to [31].

In quantum computing, a *qubit* serves as the fundamental unit of quantum information, analogous to a classical bit. A *pure quantum state* in a quantum system comprising n qubits, is represented by a unit-length vector in the 2^n -dimensional Hilbert space. A commonly used basis of the space is the *computational basis* $\{|i\rangle = |i_1, i_2, \dots, i_n\rangle : i_1, \dots, i_n \in \{0, 1\}\}$. Then any pure state $|\psi\rangle$ can be expressed as $\sum_{i=0}^{2^n-1} \alpha_i |i\rangle$, where α_i are complex numbers known as *amplitudes*, satisfying the condition $\sum_i |\alpha_i|^2 = 1$. A *mixed quantum state*, also known as a *density matrix*, represents a probability mixture of pure states. If a quantum system is in state $|\psi_i\rangle$ with probability p_i , then its density matrix $\rho := \sum_i p_i |\psi_i\rangle \langle \psi_i|$ where $\langle \psi_i|$ denotes the conjugate transpose of $|\psi_i\rangle$. Any density matrix is Hermitian and trace one. In this paper, we also use density matrix to describe classical probability distribution: If a random variable X takes value x_i with probability p_i , then it can be described by the density matrix $\sum_i p_i |x_i\rangle \langle x_i|$.

Transformations in an n -qubit quantum system are described by unitary transformations in the 2^n -dimensional Hilbert space. Such a transformation is depicted by a unitary matrix U , which satisfies $UU^\dagger = \mathbb{I}$ where \dagger is conjugate transpose and \mathbb{I} is identity matrix. If U is applied to a pure state $|\psi\rangle$, the state becomes $U|\psi\rangle$. If U is applied to a mixed state ρ , the state becomes $U\rho U^\dagger$.

Another important operation is quantum measurement. We will only use projective measurement in our paper. A projective measurement \mathcal{M} is described by a collection of orthogonal projectors $\{\Pi_i\}$ such that $\sum_i \Pi_i = \mathbb{I}$. When \mathcal{M} is applied on a pure state $|\varphi\rangle$, it collapses to state $\frac{1}{\sqrt{\beta}} \Pi_i |\varphi\rangle$ with probability $\beta = \langle \varphi | \Pi_i | \varphi \rangle$. In the language of density matrix, we have $\mathcal{M}(\rho) = \sum_i \Pi_i \rho \Pi_i$. In particular, the *computational basis measurement* has projectors $\{|i\rangle \langle i| : 0 \leq i < 2^n\}$. If a quantum state $\sum_i \alpha_i |i\rangle$ is measured in computational

basis, it collapses to state $|i\rangle$ with probability $|\alpha_i|^2$. Measurement can also be conducted on a portion of the system or on select qubits within the system. For instance, the measurement restricted on the first qubit of a n -qubit system has projectors $\{|0\rangle\langle 0| \otimes \mathbb{I}_{2^{n-1}}, |1\rangle\langle 1| \otimes \mathbb{I}_{2^{n-1}}\}$ where $\mathbb{I}_{2^{n-1}}$ is the identity operator on the last $n - 1$ qubits.

3.2 Byzantine Agreement Problem

In a Byzantine agreement problem, n distinct players labeled from 1 to n need to reach a decision on the value of a bit. Each player i inputs a bit $x_i \in \{0, 1\}$ and must decide an output bit in $\{0, 1\}$ that satisfies the following conditions:

1. **Agreement:** All uncorrupted players decide the same value.
2. **Validity:** If all x_i are the same bit y , then all uncorrupted players decide y .
3. **Termination:** All uncorrupted players terminate with probability 1.

The problem was introduced by Pease, Shostak and Lamport [32] in 1980. One can consider different network models, models of inter-player communication, models of local computation, and fault models. In this paper, the following models are of interest.

- **Network Models.** We will consider both synchronous network, where all messages are guaranteed to be delivered within some known time Δ from when they are sent, and asynchronous network where messages may be arbitrarily delayed.
- **Models of Inter-player Communications.** Every two players are connected by a transmit reliable⁴ channel. We consider two different communication paradigms: classical and quantum. In the classical model, players can communicate classical messages, while in the quantum model, they can communicate quantum messages.⁵
- **Models of Local Computation.** In the field of Byzantine protocols, there is a common tendency to overlook the intricacies of local computations. We assume players have unbounded computational power and local memory.
- **Fault models.** We model the faulty behavior of the system by an *adversary*. The adversary can corrupt participating players and make them deviate from their prescribed programs. Once a player has been corrupted, it remains corrupted permanently. The uncorrupted players are referred to as “good” and sometimes the corrupted players are labeled as “bad”. In our work, we consider the following types of adversarial behavior:
 - **Adaptive.** We will consider *adaptive* adversaries in this paper. An adaptive adversary corrupts players dynamically based on its current information at any time of the protocol.
 - **Unbound Computation.** Just like good players, the adversary has unlimited computational power and memory.
 - **Private-channel and Full-information.** We will consider both private-channel and full-information adversaries. An adversary in the private-channel model is characterized by its lack of adaptation based on the specific contents of messages exchanged within a system. Essentially, this type of adversary can only discern patterns of communication, such as the timing and players involved in message exchanges, without access to the actual message contents. By contrast, a full-information adversary possesses comprehensive knowledge of all local variables associated with the players involved in the system. In the context of the quantum model, a full information adversary knows at each point the exact pure state of the system.

⁴ Messages will not be corrupted or lost during transmission.

⁵ Since classical messages can also be encoded by qubits, no additional classical channels are required.

- **Fail-stop and Byzantine.** We will consider both Fail-stop and Byzantine adversaries. The players corrupted by the Fail-stop adversary will no longer take part in the protocol. We remark that a private-channel Fail-stop adversary *cannot* read the local memory of corrupted players.⁶ However, the players corrupted by a Byzantine adversary can deviate arbitrarily from the protocol.

We are interested in several metrics that measure the performance of BA protocols:

- **Resilience:** the maximum number of parties that can be corrupted within the protocol.
- **Round Complexity:** Assume there is a virtual “global clock” within the network that is not accessible to any player. In this context, the term *delay* refers to the time taken from sending a message to its reception. The *number of rounds*⁷ in an execution refers to the total execution time divided by the longest message delay. The *round complexity* of a protocol \mathcal{P} is defined as the maximum expected number of rounds in \mathcal{P} ’s executions, considering all inputs and potential adversaries.
- **Communication Complexity:** the maximum expected number of messages sent by good players throughout the protocol, considering all inputs and potential adversaries.

3.3 Helper lemmas

The following two lemmas will be used, of which the proofs are given in Appendix A.

► **Lemma 1.** *Let \mathcal{M} be the computational basis measurement of a Hilbert space \mathcal{H} . Then \mathcal{M} commutes with*

1. *any permutation unitary U acting on \mathcal{H} ;*
2. *any orthogonal projector Π on \mathcal{H} in computational basis.*

► **Lemma 2.** *Let G be good players’ registers, B be bad players’ registers. Initially G and B are independent and then they make quantum communication for several rounds. Assume G keeps a local copy of the communication transcript between G and B . Then the pure state of the system GB can be written as $\sum_m \alpha_m |m, \phi_m\rangle_G \otimes |\psi_m\rangle_B$ where $|m\rangle$ are the communication transcripts, $|\phi_m\rangle$ are states of G besides the communication transcripts, and $|\psi_m\rangle$ are states of B .*

4 Proof of Main Theorem

In this section, we prove our main theorem by giving a general reduction from quantum full-information BA protocols to classical private-channel protocols.

► **Theorem 1.** *Given a classical synchronous (resp. asynchronous) non-erasing BA protocol designed to counter a private-channel Fail-stop (resp. Byzantine) adversary, we can construct a quantum synchronous (resp. asynchronous) BA protocol capable of handling a full-information Fail-stop (resp. Byzantine) adversary while maintaining the same levels of resilience, round complexity, and communication complexity.*

Our reduction requires a “non-erasing” property of classical private-channel protocols:

⁶ Some BA protocols consider a stronger Fail-stop adversary who can read the memory of corrupted players, but our Theorem 1 still applies to those protocols because we only require security against a weaker Fail-stop adversary.

⁷ In the synchronous model, this definition is equivalent to the number of synchronous rounds during the execution.

► **Definition 3** (Non-erasing BA protocol). *In the context of a classical BA protocol denoted as \mathcal{P} , each computational step performed by a player can be seen as the evaluation of a function $f(s)$ where s is the internal state of the player. Consider a modified protocol, denoted as \mathcal{P}' , which follows the structure of \mathcal{P} except that players in \mathcal{P}' keep a copy of their previous state s in their local memory subsequent to each evaluation of $f(s)$.*

A BA protocol such as \mathcal{P} is called non-erasing if the adjusted protocol \mathcal{P}' maintains the characteristics of being a BA protocol while preserving the same level of resilience, round and communication complexity as \mathcal{P} .

To the best of our knowledge, this non-erasing property is considered a reasonable assumption as it is met by all existing protocols within the scope of information-theoretic BA with probability one, e.g., [16, 19, 3, 4]. Beyond our scope, there exist BA protocols requiring the ability to securely erase intermediate secrets, often referred to as the *memory-erasure model* [17]. Those protocols either rely on cryptographic assumptions [13] or succeed only with high probability [28].

The rest of this section is to prove Theorem 1. For simplicity, we will only give a full proof for the synchronous model (Section 4.1) and then briefly discuss how to extend it to the asynchronous case (Section 4.2).

4.1 Synchronous Model

In this subsection, we prove Theorem 1 for the synchronous model. Without loss of generality, we assume a synchronous classical non-erasing private-channel BA protocol \mathcal{P}_C has the following normal form.

Classical protocol \mathcal{P}_C . Let k denote the round number, $m_k^{(i,j)}$ denote the message sent from i to j and $m_k'^{(i,j)}$ denote a copy of $m_k^{(i,j)}$ to be kept by i , $b_k^{(i,j)} \in \{0, 1\}$ denote the message pattern which is 1 if $m_k^{(i,j)}$ is non-empty, and $d_k^{(i)} \in \{0, 1, \perp\}$ denote the decided value of i (\perp if not decided yet). We also use $m_k^{(*,i)}$ to denote the vector $(m_k^{(1,i)}, m_k^{(2,i)}, \dots, m_k^{(n,i)})$ and $m_k'^{(i,*)}, m_k^{(i,*)}, b_k^{(i,*)}$ are defined similarly. At round k , player i on input x_i executes the following steps.

\mathcal{P}_C for player i at round k

1. Receive messages $m_{k-1}^{(*,i)}$ from other players if $k > 1$.
2. Sample randomness $r_k^{(i)}$.
3. Compute a function $f_P : \text{View}_k^{(i)} \rightarrow (m_k^{(i,*)}, m_k'^{(i,*)}, b_k^{(i,*)}, d_k^{(i)})$ where^a

$$\text{View}_k^{(i)} := \begin{cases} (i, x_i, r_1^{(i)}) & \text{if } k = 1 \\ (\text{View}_{k-1}^{(i)}, m_{k-1}'^{(i,*)}, m_{k-1}^{(*,i)}, r_k^{(i)}) & \text{otherwise} \end{cases} .$$

4. If the decided value $d_k^{(i)} \neq \perp$, output value $d_k^{(i)}$ and terminate.^b
5. For $j \in [n]$, send messages $m_k^{(i,j)}$ to player j if $b_k^{(i,j)} = 1$.

^a Keeping $\text{View}_k^{(i)}$ in memory does not lose generality because \mathcal{P}_C is non-erasing.

^b We assume a player decides and terminates at the same time, since otherwise we can always defer the decision until the player terminates.

Then we construct a quantum BA protocol \mathcal{P}_Q by quantizing \mathcal{P}_C as follows. The essential idea is to purify the local randomness, compute everything reversibly, and do as little measurement as possible. In this way, only a superposition of all possible local information is revealed to the quantum full-information adversary. Formally,

Quantum protocol \mathcal{P}_Q . Let k denote the round number, $M_k^{(i,j)}$, $M_k^{\prime(i,j)}$, $B_k^{(i,j)}$, $D_k^{(i)}$, $R_k^{(i)}$ denote the quantum registers holding the message from player i to player j , the copy of the message, the message pattern, the decided value of player i , and the randomness of player i respectively. At round k , player i on input x_i executes the following steps.

\mathcal{P}_Q for player i at round k

1. Receive quantum messages $M_{k-1}^{(*,i)}$ from other players if $k > 1$.
2. Prepare a quantum state $\sum_r \sqrt{\Pr[r_k^{(i)} = r]} |r\rangle$ in a new quantum register $R_k^{(i)}$.
3. Let $U_P^{(i)}$ denote the unitary $|v\rangle |y\rangle \rightarrow |v\rangle |y + f_P(v)\rangle$ which reversibly computes function f_P . Execute U_P on register $\text{View}_k^{(i)}$ and an empty ancilla register $A_k^{(i)} := (M_k^{(i,*)}, M_k^{\prime(i,*)}, B_k^{(i,*)}, D_k^{(i)})$ where

$$\text{View}_k^{(i)} := \begin{cases} |i\rangle \langle i| \otimes |x_i\rangle \langle x_i| \otimes R_1^{(i)} & \text{if } k = 1 \\ (\text{View}_{k-1}^{(i)}, M_{k-1}^{\prime(i,*)}, M_{k-1}^{(*,i)}, R_k^{(i)}) & \text{otherwise} \end{cases}.$$

4. Measure register $D_k^{(i)}$. If the result $d_k^{(i)} \neq \perp$, output $d_k^{(i)}$ and terminate.
5. For each $j \in [n]$, measure $B_k^{(i,j)}$. If the result $b_k^{(i,j)} = 1$, send the $M_k^{(i,j)}$ to player j .

In the rest of this subsection, for both Fail-stop and Byzantine cases, we prove that \mathcal{P}_Q is a quantum full-information BA protocol with the same resilience, round and communication complexity as \mathcal{P}_C . The proof follows the argument that assuming there is quantum full-information adversary \mathcal{A}_Q attacking \mathcal{P}_Q , we can construct a classical adversary \mathcal{A}_C in the private-channel model attacking \mathcal{P}_C .

4.1.1 Fail-stop adversary

Without loss of generality, we assume the adversary launches attacks at the beginning of each round for both \mathcal{P}_C and \mathcal{P}_Q . The Fail-stop adversary has the ability to adaptively halt some players and choose only a subset of their messages in this round to be received. Now consider a quantum full-information Fail-stop adversary \mathcal{A}_Q attacking \mathcal{P}_Q , which can be formalized as follows.

Quantum full-information adversary \mathcal{A}_Q . Assume \mathcal{A}_Q samples its randomness r_A before the protocol starts. Then at round k , \mathcal{A}_Q first chooses the set of corrupted players S_k up to round k such that $|S_k| \leq t$ and $S_k \supseteq S_{k-1}$, and then \mathcal{A}_Q decides only a subset of $S_k \setminus S_{k-1}$'s messages to be sent. Here, we model the message exchanging step as a permutation unitary V_k which swaps the registers $M_k^{(i,j)}$ and the receiving register of player j for $i, j \in [n]$. Then \mathcal{A}_Q 's attack can be modeled by choosing an appropriate V_k . Thus \mathcal{A}_Q can be viewed as a function $f_A : r_A, \text{View}_1, \text{View}_2, \dots, \text{View}_{k-1} \rightarrow (S_k, V_k)$ where View_j is the pure state view of the system at round j .

Let $b_j := (b_j^{(1,1)}, \dots, b_j^{(n,n)})$ and $d_j := (d_j^{(1)}, \dots, d_j^{(n)})$. Observe that the randomness of the system comes only from classical variables $r_A, \{b_j\}, \{d_j\}$, so the pure state View_k is fully determined by those variables. Thus there exists a function f_V such that

32:12 Quantum Byzantine Agreement Against Full-Information Adversary

$f_V(r_A, b_1, d_1, b_2, d_2, \dots, b_k, d_k) = \text{View}_k$. Since the variables $\{b_j\}$ and $\{d_j\}$ in \mathcal{P}_C are also available to classical private-channel adversaries, now we construct a classical private-channel adversary \mathcal{A}_C attacking \mathcal{P}_C .

Classical adversary \mathcal{A}_C in the private-channel model. First sample the same randomness r_A as \mathcal{A}_Q before the protocol starts. Then at round k , compute its action by the following steps.

1. For each $j \in [k-1]$, compute quantum state $|\psi_j\rangle := f_V(r_A, b_1, d_1, b_2, d_2, \dots, b_j, d_j)$.
2. Compute action $(S_k, V_k) := f_A(r_A, |\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_{k-1}\rangle)$.

Then we prove that \mathcal{A}_C perfectly simulates the execution of $(\mathcal{P}_Q, \mathcal{A}_Q)$ when interacting with \mathcal{P}_C , which is characterized by Lemma 5.

► **Definition 4.** A k -round execution \mathcal{E} of $(\mathcal{P}_C, \mathcal{A}_C)$ is a sequence $r_A, (b_1, d_1), (b_2, d_2), \dots, (b_k, d_k)$. \mathcal{E} is also a k -round execution of $(\mathcal{P}_Q, \mathcal{A}_Q)$ since the pure states of the system at each round can completely determined by \mathcal{E} using f_V .

► **Lemma 5.** Any k -round execution \mathcal{E} occurs in $(\mathcal{P}_Q, \mathcal{A}_Q)$ and $(\mathcal{P}_C, \mathcal{A}_C)$ with the same probability. Furthermore, if the pure state after \mathcal{E} in $(\mathcal{P}_Q, \mathcal{A}_Q)$ is $\sum_u \alpha_u |u\rangle$, then the distribution of the system's possible states after \mathcal{E} in $(\mathcal{P}_C, \mathcal{A}_C)$ conditioned on \mathcal{E} is $\sum_u |\alpha_u|^2 |u\rangle \langle u|$.⁸

Proof. See Appendix B. ◀

By the above lemma, we have:

► **Proposition 6.** In the synchronous Fail-stop model, given a non-erasing classical private-channel BA protocol \mathcal{P}_C , there exists a quantum full-information BA protocol \mathcal{P}_Q with the same resilience, round and communication complexity as \mathcal{P}_C .

Proof. Assuming there exists an adversary \mathcal{A}_Q that can cause an inconsistent, invalid or non-terminating execution \mathcal{E} in $(\mathcal{P}_Q, \mathcal{A}_Q)$ with probability $p > 0$ by corrupting $\leq t$ players, then \mathcal{E} also occurs in $(\mathcal{P}_C, \mathcal{A}_C)$ with probability p by Lemma 5, which gives a contradiction. Thus the resilience of \mathcal{P}_Q is at least the resilience of \mathcal{P}_C .

Given an execution \mathcal{E} , let $|\mathcal{E}|$ be the number of rounds of \mathcal{E} , and $\text{CC}(\mathcal{E}) := \sum_{k=1}^{|\mathcal{E}|} \sum_{i \in \bar{S}_k, j \in [n]} b_k^{(i,j)}$ denote the number of messages in \mathcal{E} . Then by Lemma 5,⁹

$$\begin{aligned} \text{RC}(\mathcal{P}_Q) &:= \max_{\mathcal{A}_Q \text{ execution } \mathcal{E}} \mathbb{E} \Pr[\mathcal{E} \in (\mathcal{P}_Q, \mathcal{A}_Q)] \cdot |\mathcal{E}| \\ &= \max_{\mathcal{A}_C \in \mathcal{Q} \text{ execution } \mathcal{E}} \mathbb{E} \Pr[\mathcal{E} \in (\mathcal{P}_C, \mathcal{A}_C)] \cdot |\mathcal{E}| \leq \text{RC}(\mathcal{P}_C), \\ \text{CC}(\mathcal{P}_Q) &:= \max_{\mathcal{A}_Q \text{ execution } \mathcal{E}} \mathbb{E} \Pr[\mathcal{E} \in (\mathcal{P}_Q, \mathcal{A}_Q)] \cdot \text{CC}(\mathcal{E}) \\ &= \max_{\mathcal{A}_C \in \mathcal{Q} \text{ execution } \mathcal{E}} \mathbb{E} \Pr[\mathcal{E} \in (\mathcal{P}_C, \mathcal{A}_C)] \cdot \text{CC}(\mathcal{E}) \leq \text{CC}(\mathcal{P}_C) \end{aligned}$$

where $\text{RC}(\cdot)$ denotes round complexity, $\text{CC}(\cdot)$ denotes communication complexity, and \mathcal{Q} denotes the set of classical private-channel adversaries that are constructed from some quantum full-information adversary in the beyond way. ◀

⁸ We use density matrix to represent classical probability distribution. See Section 3.1 for details.

⁹ For simplicity, we can assume players' input of is chosen by the adversary, so there is no need to take maximum over the input.

4.1.2 Byzantine adversary

For the Byzantine case, we also assume the adversary launches attacks at the beginning of each round. Unlike the Fail-stop adversary, the Byzantine adversary can manipulate corrupted players in an arbitrary way. Now consider a quantum full-information Byzantine adversary \mathcal{A}_Q attacking \mathcal{P}_Q , which can be formalized as follows.

Quantum full-information adversary \mathcal{A}_Q . Assume \mathcal{A}_Q samples its randomness r_A before the protocol starts. Let S_k denote the corrupted players up to round k such that $|S_k| \leq t$, $S_k \supseteq S_{k-1}$, and $\bar{S}_k := [n] \setminus S_k$ denote good players. Here, we model the message-exchanging step differently from the Fail-stop case. When player j receives the message from i , the register $M_k^{(i,j)}$ is simply appended to j 's workspace. Then at round k , \mathcal{A}_Q acts as follows.

1. First let current corrupted players S_{k-1} receive all the messages sent to them.
2. Apply arbitrary quantum operation on S_{k-1} , which can be decomposed as a unitary U_k and a measurement operator \mathcal{M}_k on the registers of S_{k-1} by Stinespring dilation theorem.¹⁰ Let a_k denote the measurement outcome.
3. Choose an enlarged set S_k of corrupted players and corrupt $S_k \setminus S_{k-1}$.
4. Apply arbitrary quantum operation on S_k , which can be decomposed as applying a unitary U'_k and a measurement operator \mathcal{M}'_k on the registers of S_k . Let a'_k denote the measurement outcome.

We remark that step 4 is necessary because an adaptive adversary can decide to corrupt a player i and stop (or change) the message just sent by i in step 5 of the previous round.

Similar to the Fail-stop case, the adversary's operations $U_k, \mathcal{M}_k, U'_k, \mathcal{M}'_k$ and the corrupted set S_k are all functions of randomness r_A and the system's pure states at each step. And the system's pure states can be fully determined by classical variables $r_A, \{a_j\}, \{a'_j\}, \{b_j\}$ and $\{d_j\}$. Thus we can define two functions g_A and f_A such that

$$g_A(r_A, a_1, a'_1, b_1, d_1, \dots, a_{k-1}, a'_{k-1}, b_{k-1}, d_{k-1}) = (U_k, \mathcal{M}_k), \text{ and}$$

$$f_A(r_A, a_1, a'_1, b_1, d_1, \dots, a_{k-1}, a'_{k-1}, b_{k-1}, d_{k-1}, a_k) = (S_k, U'_k, \mathcal{M}'_k).$$

Additionally, we define $\Phi(r_A, a_1, a'_1, b_1, d_1, \dots, a_{k-1}, a'_{k-1}, b_{k-1}, d_{k-1}, a_k)$ to be the system's pure state right after step 3 of \mathcal{A}_Q at round k . Then by Lemma 2, we have

$$\Phi(r_A, a_1, a'_1, b_1, d_1, \dots, a_{k-1}, a'_{k-1}, b_{k-1}, d_{k-1}, a_k) = \sum_m \alpha_m |m, \phi_m\rangle_{\bar{S}_k} |\psi_m\rangle_{S_k} \quad (1)$$

where $|m\rangle$ are the copy of messages between \bar{S}_k and S_k kept by \bar{S}_k , $|\phi_m\rangle$ are states of \bar{S}_k besides the copy, and $|\psi_m\rangle$ are states of S_k .

Since classical variables $\{b_k\}, \{d_k\}$ in \mathcal{P}_C are also available to the adversary in the private-channel model, we can construct a classical Byzantine adversary \mathcal{A}_C attacking \mathcal{P}_C as follows.

¹⁰Stinespring dilation theorem [14] states that for any quantum operation \mathcal{E} , there exists a unitary U and an environment space E such that $\mathcal{E}(\rho) = \text{Tr}_E(U(\rho \otimes |0\rangle\langle 0|_E)U^\dagger)$. The partial trace Tr_E is equivalent to measuring E . We can assume players start with large enough empty workspace so there is no need to append new ancilla space in order to perform U .

Classical adversary \mathcal{A}_C in the private-channel model. First sample the same randomness r_A as \mathcal{A}_Q before the protocol starts. During the protocol, \mathcal{A}_C maintains a communication transcript T between good players \bar{S}_k and bad players S_k . Also, \mathcal{A}_C classically simulates a quantum state of the registers of S_k , which is denoted by $|\varphi_k\rangle$ after round k . At round k , \mathcal{A}_C acts as follows.

1. Let S_{k-1} receive all the messages $m_{k-1}^{(*,S_{k-1})}$ sent to them and record in T .
2. Compute (U_k, \mathcal{M}_k) by g_A . Then apply U_k and \mathcal{M}_k on $|\varphi_{k-1}\rangle \otimes |m_{k-1}^{(*,S_{k-1})}\rangle$ and obtain the measurement outcome a_k .
3. Compute $(S_k, U'_k, \mathcal{M}'_k)$ by f_A . Corrupt players S_k and update T as the communication transcript between new sets \bar{S}_k and S_k . Then according to T , \mathcal{A}_C discards old state $|\varphi_{k-1}\rangle$ and simulates a new state $|\psi_T\rangle$ which is defined in Eq. (1).
4. Apply U'_k and \mathcal{M}'_k to $|\psi_T\rangle$ and obtain measurement outcome a'_k . Then apply a computational basis measurement \mathcal{M}_{msg} on messages to be sent from S_k to \bar{S}_k and add those messages to T . Let $|\varphi_k\rangle$ be the pure state after applying U'_k , \mathcal{M}'_k and \mathcal{M}_{msg} .

► **Definition 7.** A k -round execution \mathcal{E} of $(\mathcal{P}_C, \mathcal{A}_C)$ is a sequence $r_A, (a_1, a'_1, b_1, d_1), \dots, (a_k, a'_k, b_k, d_k)$. \mathcal{E} is also a k -round execution of $(\mathcal{P}_Q, \mathcal{A}_Q)$ since the pure states of the system can be determined by \mathcal{E} .

► **Lemma 8.** Any k -round execution \mathcal{E} occurs in $(\mathcal{P}_Q, \mathcal{A}_Q)$ and $(\mathcal{P}_C, \mathcal{A}_C)$ with the same probability. Furthermore, if the pure state in $(\mathcal{P}_Q, \mathcal{A}_Q)$ after \mathcal{E} is $|Q_k\rangle$, then the distribution of system's state in $(\mathcal{P}_C, \mathcal{A}_C)$ after \mathcal{E} is $C_k := \mathcal{M}_{\bar{S}_k}(|Q_k\rangle\langle Q_k|)$ where $\mathcal{M}_{\bar{S}_k}$ is the computational basis measurement on good players \bar{S}_k 's registers.¹¹

Proof. See Appendix C. ◀

By the above lemma, we conclude Theorem 1 for the synchronous Byzantine case, which can be proven the same way as the Fail-stop case (Proposition 6).

► **Proposition 9.** In the synchronous Byzantine model, given a non-erasing classical private-channel BA protocol \mathcal{P}_C , there exists a quantum full-information BA protocol \mathcal{P}_Q with the same resilience, round and communication complexity as \mathcal{P}_C .

4.2 Asynchronous Model

The techniques used in the proof above can be extended to the asynchronous model as well. However, a key distinction lies in the terminology used to characterize the execution: while the synchronous model employs “rounds”, the asynchronous model employs “steps”. In this context, a step involves a single good player receiving only one message, carrying out computations, and potentially transmitting messages. The order in which players receive messages is determined by the adversary. For simplicity, we assume that each player initially receives its input as its first message, and each message contains the sender’s ID.

In alignment with the synchronous model, our approach involves first giving a normal form to any asynchronous classical non-erasing private-channel BA protocol \mathcal{P}_C and then quantizing it into a quantum protocol \mathcal{P}_Q against the full-information adversary.

¹¹Density matrix C_k represents a distribution of system’s states with classical \bar{S}_k and quantum S_k . It is classically feasible because S_k ’s quantum state is classically simulated, and the correlation between \bar{S}_k and S_k is classical, i.e., there is no quantum entanglement.

Classical protocol \mathcal{P}_C . Each player is activated each time it receives a message. Let $\pi_k^{(i)}$ denote the k -th message player i receives, where the first message $\pi_1^{(i)}$ is its input x_i . The notations $m_k^{(i,j)}, m_k^{\prime(i,j)}, b_k^{(i)}, d_k^{(i)}$ are defined similarly as in synchronous model (Section 4.1).

\mathcal{P}_C for player i upon receiving the k -th message $\pi_k^{(i)}$

1. Sample randomness $r_k^{(i)}$.
2. Compute a function $f_P : \text{View}_k^{(i)} \rightarrow (m_k^{(i,*)}, m_k^{\prime(i,*)}, b_k^{(i,*)}, d_k^{(i)})$ where

$$\text{View}_k^{(i)} := \begin{cases} (i, x_i, r_1^{(i)}) & \text{if } k = 1 \\ (\text{View}_{k-1}^{(i)}, m_{k-1}^{\prime(i,*)}, \pi_k^{(i)}, r_k^{(i)}) & \text{otherwise} \end{cases}.$$

3. If the decided value $d_k^{(i)} \neq \perp$, output value $d_k^{(i)}$ and terminate.
4. For $j \in [n]$, send messages $m_k^{(i,j)}$ to player j if $b_k^{(i,j)} = 1$.

Quantum protocol \mathcal{P}_Q . Each player is activated each time it receives a message. Let $\Pi_k^{(i)}$ denote the k -th quantum message player i receives. The notations $M_k^{(i,j)}, M_k^{\prime(i,j)}, B_k^{(i,j)}, D_k^{(i)}$ are defined similarly as in synchronous model (Section 4.1). We remark that $\Pi_k^{(i)}$ is an alias of register $M_{k'}^{(j',i)}$ for some j', k' .

\mathcal{P}_Q for player i upon receiving the k -th message $\Pi_k^{(i)}$

1. Prepare a quantum state $\sum_r \sqrt{\Pr[r_k^{(i)} = r]} |r\rangle$ in a new quantum register $R_k^{(i)}$.
2. Let $U_P^{(i)}$ denote the unitary $|v\rangle |y\rangle \rightarrow |v\rangle |y + f_P(v)\rangle$ which reversibly computes function f_P . Execute $U_P^{(i)}$ on register $\text{View}_k^{(i)}$ and an empty ancilla register $A_k^{(i)} := (M_k^{(i,*)}, M_k^{\prime(i,*)}, B_k^{(i,*)}, D_k^{(i)})$ where

$$\text{View}_k^{(i)} := \begin{cases} |i\rangle \langle i| \otimes |x_i\rangle \langle x_i| \otimes R_1^{(i)} & \text{if } k = 1 \\ (\text{View}_{k-1}^{(i)}, M_{k-1}^{\prime(i,*)}, \Pi_k^{(i)}, R_k^{(i)}) & \text{otherwise} \end{cases}.$$

3. Measure register $D_k^{(i)}$. If the result $d_k^{(i)} \neq \perp$, output $d_k^{(i)}$ and terminate.
4. For each $j \in [n]$, measure $B_k^{(i,j)}$. If the result $b_k^{(i,j)} = 1$, send the $M_k^{(i,j)}$ to player j .

Then we claim that \mathcal{P}_Q is a quantum BA protocol against the quantum full-information adversary in the asynchronous model with the same round and communication complexity as \mathcal{P}_C . The proof is almost the same as the synchronous case, so we only sketch the proof here.

Assuming there is a quantum full-information Fail-stop (resp. Byzantine) adversary \mathcal{A}_Q attacking \mathcal{P}_Q , we can construct a classical private-channel Fail-stop (resp. Byzantine) adversary \mathcal{A}_C attacking \mathcal{P}_C as in Section 4.1.1 (resp. Section 4.1.2). Then we can define execution execution in the asynchronous model.

► **Definition 10 (Informal).** A k -step execution \mathcal{E} is defined to be a sequence $r_A, (a_1, b_1, d_1), (a_2, b_2, d_2), \dots, (a_k, b_k, d_k)$ where r_A is the adversary's randomness, a_j is some classical information the adversary obtains at step j , $b_j \in \{0, 1\}^n$ is the message pattern, and $d_j \in \{0, 1, \perp\}$ is the decided value of the player activated at step j .

Then similar to Lemma 5 (resp. Lemma 8), we prove that any execution occurs in $(\mathcal{P}_Q, \mathcal{A}_Q)$ with the same probability.

► **Lemma 11** (Informal). *Any k -step execution \mathcal{E} occurs in $(\mathcal{P}_Q, \mathcal{A}_Q)$ and $(\mathcal{P}_C, \mathcal{A}_C)$ with the same probability.*

Since the information contained in an execution \mathcal{E} fully determines the properties, number of rounds, and number of messages of the protocol, we can conclude Theorem 1 in the asynchronous case. This can be proven similarly to the synchronous case (Proposition 6 and Proposition 9).

► **Proposition 12.** *In the asynchronous Fail-stop (or Byzantine) model, given a non-erasing classical private-channel BA protocol \mathcal{P}_C , there exists a quantum full-information BA protocol \mathcal{P}_Q with the same resilience, round and communication complexity as \mathcal{P}_C .*

Proof of Theorem 1. Theorem 1 can be obtained by integrating the results from Proposition 6, Proposition 9 and Proposition 12. ◀

5 Discussions

In this paper, we present a general reduction from quantum full-information BA protocols to classical private-channel BA protocols that preserves resilience, round and communication complexity. Utilizing this reduction, we make progress towards the open question posed by [7] of whether quantum BA can achieve $O(1)$ round complexity and optimal resilience $t < n/3$ simultaneously in the asynchronous full-information model. We show that $O(1)$ round complexity and suboptimal resilience $t < n/(3 + \epsilon)$ is possible for any constant $\epsilon > 0$. Our reduction also suggests that designing a better classical private-channel protocol may finally lead to the resolution of this open question.

There are several interesting directions for future research. Firstly, it would be valuable to explore whether the reverse of our reduction is possible, i.e., whether any quantum full-information BA protocol can be converted to a classical private-channel BA protocol without compromising key attributes like resilience. Existing techniques in this paper do not apply due to the ability of good players to employ quantum operations. Secondly, it is worth considering the potential generalization of our results to less strict models, such as BA that terminates only with high probability [12, 28], or BA that requires erasing intermediate states [28, 13]. Thirdly, it is worthwhile to explore the potential for developing BA protocols with improved performance by granting quantum players the ability to utilize private memory, thereby shifting the adversary from a position of full-information to one of limited knowledge. This model presents an intriguing opportunity for innovation, especially considering the existence of quantum key distribution in such a framework [9]. Finally, while our primary focus is on addressing the BA problem as it stands as a fundamental challenge in this field, we anticipate that our methods can also be applied to other fault-tolerant distributed computing tasks like coin toss and leader election.

References

- 1 Ittai Abraham, Danny Dolev, and Joseph Y Halpern. An almost-surely terminating polynomial protocol for asynchronous byzantine agreement with optimal resilience. In *Proceedings of the Twenty-seventh ACM symposium on Principles of Distributed Computing*, pages 405–414, 2008. doi:10.1145/1400751.1400804.
- 2 Hagit Attiya and Keren Censor. Tight bounds for asynchronous randomized consensus. *Journal of the ACM (JACM)*, 55(5):1–26, 2008. doi:10.1145/1411509.1411510.
- 3 Hagit Attiya and Jennifer Welch. *Distributed Computing: Fundamentals, Simulations and Advanced Topics*. John Wiley & Sons, Inc., Hoboken, NJ, USA, 2004.

- 4 Laasya Bangalore, Ashish Choudhury, and Arpita Patra. Almost-surely terminating asynchronous byzantine agreement revisited. In *Proceedings of the 2018 ACM Symposium on Principles of Distributed Computing*, pages 295–304, 2018. URL: <https://dl.acm.org/citation.cfm?id=3212735>.
- 5 Ziv Bar-Joseph and Michael Ben-Or. A tight lower bound for randomized synchronous consensus. In *Proceedings of the Seventeenth Annual ACM Symposium on Principles of Distributed Computing*, PODC '98, pages 193–199, New York, NY, USA, 1998. Association for Computing Machinery. doi:10.1145/277697.277733.
- 6 Michael Ben-Or. Another advantage of free choice: Completely asynchronous agreement protocols (extended abstract). In Robert L. Probert, Nancy A. Lynch, and Nicola Santoro, editors, *Proceedings of the Second Annual ACM Symposium on Principles of Distributed Computing, Montreal, Quebec, Canada, August 17-19, 1983*, pages 27–30. ACM, 1983. doi:10.1145/800221.806707.
- 7 Michael Ben-Or and Avinatan Hassidim. Fast quantum byzantine agreement. In *Proceedings of the Thirty-Seventh Annual ACM Symposium on Theory of Computing*, STOC '05, pages 481–485, New York, NY, USA, 2005. Association for Computing Machinery. doi:10.1145/1060590.1060662.
- 8 Michael Ben-Or and Nathan Linial. Collective coin flipping, robust voting schemes and minima of banzhaf values. In *26th Annual Symposium on Foundations of Computer Science (SFCS 1985)*, pages 408–416, 1985. doi:10.1109/SFCS.1985.15.
- 9 Charles H Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. *Theoretical Computer Science*, 560:7–11, 2014. doi:10.1016/j.tcs.2014.05.025.
- 10 Gabriel Bracha. Asynchronous byzantine agreement protocols. *Information and Computation*, 75(2):130–143, 1987. doi:10.1016/0890-5401(87)90054-X.
- 11 Gabriel Bracha. An $O(\log n)$ expected rounds randomized byzantine generals protocol. *J. ACM*, 34(4):910–920, October 1987. doi:10.1145/31846.42229.
- 12 Ran Canetti and Tal Rabin. Fast asynchronous byzantine agreement with optimal resilience. In S. Rao Kosaraju, David S. Johnson, and Alok Aggarwal, editors, *Proceedings of the Twenty-fifth Annual ACM Symposium on Theory of Computing*, pages 42–51, 1993. doi:10.1145/167088.167105.
- 13 Jing Chen and Silvio Micali. Algorand: A secure and efficient distributed ledger. *Theor. Comput. Sci.*, 777:155–183, 2019. doi:10.1016/j.tcs.2019.02.001.
- 14 Man-Duen Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications*, 10(3):285–290, 1975.
- 15 Vicent Cholvi. Quantum byzantine agreement for any number of dishonest parties. *Quantum Information Processing*, 21(4):151, 2022. doi:10.1007/s11128-022-03492-y.
- 16 Benny Chor, Michael Merritt, and David B Shmoys. Simple constant-time consensus protocols in realistic failure models. *Journal of the ACM (JACM)*, 36(3):591–614, 1989. doi:10.1145/65950.65956.
- 17 Thaddeus Dryja, Quanquan C Liu, and Neha Narula. A lower bound for byzantine agreement and consensus for adaptive adversaries using vdfs. *arXiv preprint arXiv:2004.01939*, abs/2004.01939, 2020. arXiv:2004.01939, doi:10.48550/arXiv.2004.01939.
- 18 U. Feige. Noncryptographic selection protocols. In *40th Annual Symposium on Foundations of Computer Science (Cat. No.99CB37039)*, pages 142–152, 1999. doi:10.1109/SFFCS.1999.814586.
- 19 Peaseh Feldman and Silvio Micali. An optimal probabilistic protocol for synchronous byzantine agreement. *SIAM Journal on Computing*, 26(4):873–933, 1997. doi:10.1137/S0097539790187084.
- 20 Matthias Fitzi, Nicolas Gisin, and Ueli Maurer. Quantum solution to the byzantine agreement problem. *Phys. Rev. Lett.*, 87:217901, November 2001. doi:10.1103/PhysRevLett.87.217901.
- 21 Matthias Fitzi, Daniel Gottesman, Martin Hirt, Thomas Holenstein, and Adam Smith. Detectable byzantine agreement secure against faulty majorities. In *Proceedings of the Twenty-First*

- Annual Symposium on Principles of Distributed Computing*, PODC '02, pages 118–126, New York, NY, USA, 2002. Association for Computing Machinery. doi:10.1145/571825.571841.
- 22 Sascha Gaertner, Mohamed Bourennane, Christian Kurtsiefer, Adán Cabello, and Harald Weinfurter. Experimental demonstration of a quantum protocol for byzantine agreement and liar detection. *Phys. Rev. Lett.*, 100:070504, February 2008. doi:10.1103/PhysRevLett.100.070504.
 - 23 Oded Goldreich, Shafi Goldwasser, and Nathan Linial. Fault-tolerant computation in the full information model. *SIAM Journal on Computing*, 27(2):506–544, 1998. doi:10.1137/S0097539793246689.
 - 24 Mohammadtaghi Hajiaghayi, Dariusz Rafal Kowalski, and Jan Olkowski. Brief announcement: Improved consensus in quantum networks. In *Proceedings of the 2023 ACM Symposium on Principles of Distributed Computing*, pages 286–289, 2023. doi:10.1145/3583668.3594600.
 - 25 Shang-En Huang, Seth Pettie, and Leqi Zhu. Byzantine agreement in polynomial time with near-optimal resilience. In Stefano Leonardi and Anupam Gupta, editors, *STOC '22: 54th Annual ACM SIGACT Symposium on Theory of Computing, Rome, Italy, June 20 - 24, 2022*, STOC 2022, pages 502–514, New York, NY, USA, 2022. ACM. doi:10.1145/3519935.3520015.
 - 26 Shang-En Huang, Seth Pettie, and Leqi Zhu. Byzantine agreement with optimal resilience via statistical fraud detection. *Journal of the ACM*, 71(2):12:1–12:37, 2024. doi:10.1145/3639454.
 - 27 Bruce M. Kapron, David Kempe, Valerie King, Jared Saia, and Vishal Sanwalani. Fast asynchronous byzantine agreement and leader election with full information. *ACM Trans. Algorithms*, 6(4), September 2010. doi:10.1145/1824777.1824788.
 - 28 Valerie King and Jared Saia. Breaking the $O(n^2)$ bit barrier: scalable byzantine agreement with an adaptive adversary. *Journal of the ACM (JACM)*, 58(4):1–24, 2011. doi:10.1145/1989727.1989732.
 - 29 Valerie King and Jared Saia. Byzantine agreement in expected polynomial time. *J. ACM*, 63(2), March 2016. doi:10.1145/2837019.
 - 30 Qing-bin Luo, Kai-yuan Feng, and Ming-hui Zheng. Quantum multi-valued byzantine agreement based on d-dimensional entangled states. *International Journal of Theoretical Physics*, 58(12):4025–4032, 2019. doi:10.1007/s10773-019-04269-3.
 - 31 Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information (10th Anniversary edition)*. Cambridge University Press, 2016. URL: <https://www.cambridge.org/de/academic/subjects/physics/quantum-physics-quantum-information-and-quantum-computation/quantum-computation-and-quantum-information-10th-anniversary-edition?format=HB>.
 - 32 Marshall C. Pease, Robert E. Shostak, and Leslie Lamport. Reaching agreement in the presence of faults. *J. ACM*, 27(2):228–234, April 1980. doi:10.1145/322186.322188.
 - 33 Ramij Rahaman, Marcin Wieśniak, and Marek Żukowski. Quantum byzantine agreement via hardy correlations and entanglement swapping. *Phys. Rev. A*, 92:042302, October 2015. doi:10.1103/PhysRevA.92.042302.
 - 34 Alexander Russell and David Zuckerman. Perfect information leader election in $\log^*n + O(1)$ rounds. In *39th Annual Symposium on Foundations of Computer Science, FOCS '98, November 8-11, 1998, Palo Alto, California, USA*, pages 576–583. IEEE Computer Society, 1998. doi:10.1109/SFCS.1998.743508.
 - 35 Sam Toueg. Randomized byzantine agreements. In *Proceedings of the Third Annual ACM Symposium on Principles of Distributed Computing*, PODC '84, pages 163–178, New York, NY, USA, 1984. Association for Computing Machinery. doi:10.1145/800222.806744.

A Proofs of helper lemmas

A.1 Proof of Lemma 1

Proof.

1. Since U is a permutation unitary, there exists a permutation π such that $U|i\rangle = |\pi(i)\rangle$ for any computational basis $|i\rangle$ in \mathcal{H} . Given any density matrix $\rho = \sum_{i,j} \rho_{i,j} |i\rangle\langle j|$ in \mathcal{H} , we have

$$\begin{aligned} \mathcal{M}(U\rho U^\dagger) &= \mathcal{M}\left(\sum_{i,j} \rho_{i,j} |\pi(i)\rangle\langle \pi(j)|\right) \\ &= \sum_k |k\rangle\langle k| \sum_{i,j} \rho_{i,j} |\pi(i)\rangle\langle \pi(j)| |k\rangle\langle k| = \sum_k \rho_{k,k} |\pi(k)\rangle\langle \pi(k)|, \\ U\mathcal{M}(\rho)U^\dagger &= U \sum_k |k\rangle\langle k| \sum_{i,j} \rho_{i,j} |i\rangle\langle j| |k\rangle\langle k| U^\dagger \\ &= U \sum_k \rho_{k,k} |k\rangle\langle k| U^\dagger = \sum_k \rho_{k,k} |\pi(k)\rangle\langle \pi(k)|. \end{aligned}$$

Thus $\mathcal{M}(U\rho U^\dagger) = U\mathcal{M}(\rho)U^\dagger$.

2. Since Π is an orthogonal projector in the computational basis, we have $\Pi = \sum_{i \in S} |i\rangle\langle i|$ for some set S . For any state $\rho \in \mathcal{H}$, one can verify that $\mathcal{M}(\Pi\rho\Pi^\dagger) = \Pi\mathcal{M}(\rho)\Pi^\dagger$. ◀

A.2 Proof of Lemma 2

Proof. Proof by induction on the number of messages. Initially, \mathbf{G} and \mathbf{B} are independent, so the state of \mathbf{GB} is $|\phi_0\rangle_{\mathbf{G}} \otimes |\psi_0\rangle_{\mathbf{B}}$. Assuming currently the state of \mathbf{GB} is $\sum_m \alpha_m |m, \phi_m\rangle_{\mathbf{G}} \otimes |\psi_m\rangle_{\mathbf{B}}$, consider the next message. First \mathbf{G} and \mathbf{B} apply a local unitary $U_{\mathbf{G}} \otimes U_{\mathbf{B}}$ to generate messages. Note that $U_{\mathbf{G}}$ will not change the previous transcripts $|m\rangle$. Then the state becomes

$$\sum_m \alpha_m U_{\mathbf{G}} |m, \phi_m\rangle_{\mathbf{G}} \otimes U_{\mathbf{B}} |\psi_m\rangle_{\mathbf{B}} = \sum_m \alpha_m |m, \phi'_m\rangle_{\mathbf{G}} \otimes |\psi'_m\rangle_{\mathbf{B}}.$$

- If the message is sent by \mathbf{G} , then the system can be written as

$$\sum_m \alpha_m \sum_{m'} \beta_{m'} |m, m', m', \phi'_{m,m'}\rangle_{\mathbf{G}} \otimes |\psi'_m\rangle_{\mathbf{B}}$$

where the second m' is the message to be sent to \mathbf{B} and the first m' is a copy to be kept by \mathbf{G} . After sending the message, the system becomes

$$\sum_{m,m'} \alpha_m \beta_{m'} |m, m', \phi'_{m,m'}\rangle_{\mathbf{G}} \otimes |m', \psi'_m\rangle_{\mathbf{B}}.$$

- If the message is sent by \mathbf{B} , then the system can be written as

$$\sum_m \alpha_m |m, \phi_m\rangle_{\mathbf{G}} \otimes \sum_{m'} \beta_{m'} |m', \psi'_{m,m'}\rangle_{\mathbf{B}}.$$

After sending the message, the system becomes

$$\sum_{m,m'} \alpha_m \beta_{m'} |m, m', \phi_m\rangle_{\mathbf{G}} \otimes |\psi'_{m,m'}\rangle_{\mathbf{B}}. \quad \blacktriangleleft$$

B Proof of Lemma 5

Proof. Prove by induction on k . When $k = 0$, the execution $\mathcal{E}_0 = r_A$ occurs in $(\mathcal{P}_Q, \mathcal{A}_Q)$ and $(\mathcal{P}_C, \mathcal{A}_C)$ both with the probability of r_A . Assume the lemma holds for $k - 1$. Consider a k -round execution $\mathcal{E}_k := r_A, (b_1, d_1), (b_2, d_2), \dots, (b_{k-1}, d_{k-1}), (b_k, d_k)$. By inductive hypothesis, the $(k-1)$ -round prefix $\mathcal{E}_{k-1} := r_A, (b_1, d_1), (b_2, d_2), \dots, (b_{k-1}, d_{k-1})$ occurs with probability p in both $(\mathcal{P}_Q, \mathcal{A}_Q)$ and $(\mathcal{P}_C, \mathcal{A}_C)$, and the state before round k is $|Q_{k-1}\rangle := \sum_u \alpha_u |u\rangle$ and $C_{k-1} := \sum_u |\alpha_u|^2 |u\rangle \langle u|$ for $(\mathcal{P}_Q, \mathcal{A}_Q)$ and $(\mathcal{P}_C, \mathcal{A}_C)$ respectively.

Round k of $(\mathcal{P}_Q, \mathcal{A}_Q)$. After \mathcal{A}_Q 's action and players receiving messages, the state of the system becomes $V_k |Q_{k-1}\rangle$. Then good players first prepare a superposition state $|r_k\rangle$ of randomness in a new register R_k and prepare $|0\rangle$ in a new register A_k . Note that here $R_k := (R_k^{(1)}, \dots, R_k^{(n)})$, $A_k := (A_k^{(1)}, \dots, A_k^{(n)})$ and all other notations without superscript are defined similarly.

Then the players apply the unitary operator $U_P := \otimes_{i=1}^n U_P^{(i)}$ followed by a measurement which outputs (b_k, d_k) . The measurement can be viewed as an orthogonal projector Π_{b_k, d_k} in computational basis that projects the quantum state of registers (B_k, D_k) into values (b_k, d_k) . Then the state after round k becomes

$$|Q_k\rangle := \frac{1}{\sqrt{\beta}} \Pi_{b_k, d_k} U_P (V_k |Q_{k-1}\rangle \otimes |0\rangle_{A_k} \otimes |r_k\rangle_{R_k})$$

where β is the probability of getting measurement outcome (b_k, d_k) .

Round k of $(\mathcal{P}_C, \mathcal{A}_C)$. The first observation is that C_k can be viewed as first applying the same operation as $(\mathcal{P}_Q, \mathcal{A}_Q)$ and then applying computational basis measurement \mathcal{M} on the whole system:

$$C_k := \mathcal{M} \left(\frac{1}{\beta'} \Pi_{b_k, d_k} U_P V_k (C_{k-1} \otimes |0, r_k\rangle \langle 0, r_k|_{AR}) V_k^\dagger U_P^\dagger \Pi_{b_k, d_k}^\dagger \right)$$

where β' is the probability of getting measurement outcome (b_k, d_k) . The second observation is that $C_{k-1} = \mathcal{M}'(|Q_{k-1}\rangle \langle Q_{k-1}|)$ where \mathcal{M}' denotes the computational basis measurement in $|Q_{k-1}\rangle$'s space. Then

$$\begin{aligned} C_k &= \mathcal{M} \left(\frac{1}{\beta'} \Pi_{b_k, d_k} U_P V_k (\mathcal{M}'(|Q_{k-1}\rangle \langle Q_{k-1}|) \otimes |0, r_k\rangle \langle 0, r_k|_{AR}) V_k^\dagger U_P^\dagger \Pi_{b_k, d_k}^\dagger \right) \\ &= \frac{1}{\beta'} \Pi_{b_k, d_k} U_P V_k \mathcal{M} (\mathcal{M}'(|Q_{k-1}\rangle \langle Q_{k-1}|) \otimes |0, r_k\rangle \langle 0, r_k|_{AR}) V_k^\dagger U_P^\dagger \Pi_{b_k, d_k}^\dagger. \end{aligned}$$

The second equality is because U_P, V_k are all permutation unitaries and Π_{b_k, d_k} is an orthogonal projector in computational basis, which all commute with \mathcal{M} by Lemma 1. Since \mathcal{M} measures a larger space than \mathcal{M}' , \mathcal{M}' can be absorbed into \mathcal{M} , i.e., $\mathcal{M}\mathcal{M}' \equiv \mathcal{M}$. Thus

$$\begin{aligned} C_k &= \frac{1}{\beta'} \Pi_{b_k, d_k} U_P V_k \mathcal{M} (|Q_{k-1}\rangle \langle Q_{k-1}| \otimes |0, r_k\rangle \langle 0, r_k|_{AR}) V_k^\dagger U_P^\dagger \Pi_{b_k, d_k}^\dagger \\ &= \mathcal{M} \left(\frac{1}{\beta'} \Pi_{b_k, d_k} U_P V_k (|Q_{k-1}\rangle \langle Q_{k-1}| \otimes |0, r_k\rangle \langle 0, r_k|_{AR}) V_k^\dagger U_P^\dagger \Pi_{b_k, d_k}^\dagger \right) \\ &= \frac{\beta}{\beta'} \mathcal{M} (|Q_k\rangle \langle Q_k|). \end{aligned}$$

Finally, we have $\beta = \beta'$ because C_k has trace 1. Thus $C_k = \mathcal{M}(|Q_k\rangle \langle Q_k|)$ and the probability of the \mathcal{E}_k occurring is $p\beta$ for both $(\mathcal{P}_Q, \mathcal{A}_Q)$ and $(\mathcal{P}_C, \mathcal{A}_C)$. \blacktriangleleft

C Proof of Lemma 8

Proof. Prove by induction on k . The base case $k = 0$ is trivial. Assume the proposition holds for $k-1$. Consider a k -round execution $\mathcal{E}_k := r_A, (a_1, a'_1, b_1, d_1), \dots, (a_k, a'_k, b_k, d_k)$. By inductive hypothesis, the $(k-1)$ -round prefix $\mathcal{E}_{k-1} := r_A, (a_1, a'_1, b_1, d_1), \dots, (a_{k-1}, a'_{k-1}, b_{k-1}, d_{k-1})$ occurs with probability p in both $(\mathcal{P}_Q, \mathcal{A}_Q)$ and $(\mathcal{P}_C, \mathcal{A}_C)$, and the state before round k is $|Q_{k-1}\rangle$ and $C_{k-1} := \mathcal{M}_{S_{k-1}}(|Q_{k-1}\rangle\langle Q_{k-1}|)$ for $(\mathcal{P}_Q, \mathcal{A}_Q)$ and $(\mathcal{P}_C, \mathcal{A}_C)$ respectively. Since good players' messages are fully determined by their local variables, $\mathcal{M}_{S_{k-1}}$ can be restricted to a measurement $\mathcal{M}'_{S_{k-1}}$ which does not measure the messages \bar{S}_{k-1} are about to send out, i.e., $\mathcal{M}_k^{(S_{k-1}, *)}$. Then $C_{k-1} := \mathcal{M}'_{S_{k-1}}(|Q_{k-1}\rangle\langle Q_{k-1}|)$. In the following, we consider the evolution of $|Q_{k-1}\rangle$ and C_{k-1} in round k .

Step 1 and 2 of the adversary. Both \mathcal{A}_Q and \mathcal{A}_C apply U_k and \mathcal{M}_k on S_{k-1} along with the messages $\mathcal{M}_k^{(S_{k-1}, *)}$ sent to them. Since we know \mathcal{M}_k will output a_k , it can be viewed as an orthogonal projector Π_{a_k} that projects into the space of a_k . Since Π_{a_k} and U_k act only on S_{k-1} 's registers and the messages \bar{S}_{k-1} will send to S_{k-1} , they commute with $\mathcal{M}'_{S_{k-1}}$. Thus the states of $(\mathcal{P}_Q, \mathcal{A}_Q)$ and $(\mathcal{P}_C, \mathcal{A}_C)$ become

$$\begin{aligned} |Q_{k-0.5}\rangle &:= \frac{1}{\sqrt{\gamma}} \Pi_{a_k} U_k |Q_{k-1}\rangle, \text{ and} \\ C_{k-0.5} &:= \frac{1}{\gamma'} \Pi_{a_k} U_k C_{k-1} U_k^\dagger \Pi_{a_k}^\dagger = \frac{1}{\gamma'} \Pi_{a_k} U_k \mathcal{M}'_{\bar{S}_{k-1}}(|Q_{k-1}\rangle\langle Q_{k-1}|) U_k^\dagger \Pi_{a_k}^\dagger \\ &= \mathcal{M}'_{\bar{S}_{k-1}} \left(\frac{1}{\gamma'} \Pi_{a_k} U_k |Q_{k-1}\rangle\langle Q_{k-1}| U_k^\dagger \Pi_{a_k}^\dagger \right) = \frac{\gamma}{\gamma'} \mathcal{M}'_{\bar{S}_{k-1}}(|Q_{k-0.5}\rangle\langle Q_{k-0.5}|). \end{aligned}$$

where γ and γ' are probabilities of $|Q_{k-1}\rangle$ and C_{k-1} outputting a_k . Since $C_{k-0.5}$ has trace 1, we have $\gamma = \gamma'$ and thus $C_{k-0.5} = \mathcal{M}'_{\bar{S}_{k-1}}(|Q_{k-0.5}\rangle\langle Q_{k-0.5}|)$.

Step 3 of the adversary. Both \mathcal{A}_Q and \mathcal{A}_C choose an enlarged set S_k of corrupted players. This step does not affect the state $|Q_{k-0.5}\rangle$ of $(\mathcal{P}_Q, \mathcal{A}_Q)$. By Lemma 2, we have $|Q_{k-0.5}\rangle = \sum_m \alpha_m |m, \phi_m\rangle_{\bar{S}_k} |\psi_m\rangle_{S_k}$. Since $\mathcal{M}'_{\bar{S}_{k-1}}$ can be decomposed as $\mathcal{M}'_{\bar{S}_{k-1}} \equiv \mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3$, where \mathcal{M}_1 acts on transcript $|m\rangle$, \mathcal{M}_2 acts on registers of \bar{S}_{k-1} besides $|m\rangle$, and \mathcal{M}_3 acts on registers newly corrupted players $S_k \setminus S_{k-1}$, we have

$$C_{k-0.5} = \mathcal{M}'_{\bar{S}_{k-1}}(|Q_{k-0.5}\rangle\langle Q_{k-0.5}|) = \sum_m |\alpha_m|^2 |m\rangle\langle m| \otimes \mathcal{M}_2(|\phi_m\rangle\langle\phi_m|) \otimes \mathcal{M}_3(|\psi_m\rangle\langle\psi_m|).$$

In step 3 of \mathcal{A}_C , \mathcal{A}_C has recorded the transcript m and will discard the old state $\mathcal{M}_3(|\psi_m\rangle\langle\psi_m|)$ and simulate a new state $|\psi_m\rangle$. After that, the state of $(\mathcal{P}_C, \mathcal{A}_C)$ becomes

$$C'_{k-0.5} := \sum_m |\alpha_m|^2 |m\rangle\langle m| \otimes \mathcal{M}_2(|\phi_m\rangle\langle\phi_m|) \otimes |\psi_m\rangle\langle\psi_m| = \mathcal{M}'_{\bar{S}_k}(|Q_{k-0.5}\rangle\langle Q_{k-0.5}|).$$

Note that we use $\mathcal{M}'_{\bar{S}_k}$ to distinguish from operator $\mathcal{M}_{\bar{S}_k}$ which also measures the newly appended registers A_k, R_k , and $\mathcal{M}_k^{(S_k, \bar{S}_k)}$ of \bar{S}_k at round k .

Step 4 of the adversary and good players' action. Step 4 of \mathcal{A}_Q applies U'_k followed by measurement \mathcal{M}'_k which outputs a'_k on S_k 's registers. The measurement can be viewed as a projector $\Pi_{a'_k}$ that projects into the space of a'_k . Then good players apply unitary U_P and

32:22 Quantum Byzantine Agreement Against Full-Information Adversary

projector Π_{b_k, d_k} which projects the state of registers (B_k, D_k) into values (b_k, d_k) . Thus the state of $(\mathcal{P}_Q, \mathcal{A}_Q)$ after round k becomes

$$|Q_k\rangle := \frac{1}{\sqrt{\beta}} \Pi_{b_k, d_k} U_P \left(\Pi_{a'_k} U'_k |Q_{k-0.5}\rangle \otimes |0\rangle_A \otimes |r_k\rangle_R \right).$$

where β is the probability of outputting a'_k, b_k , and d_k .

Step 4 of \mathcal{A}_C will additionally apply a measurement \mathcal{M}_{msg} on messages $M_{k-1}^{(S_k, \bar{S}_k)}$ sent from S_k to \bar{S}_k . The good players' action of \mathcal{P}_C can be viewed as applying the same operation as \mathcal{P}_Q and then measuring good players \bar{S}_k 's space in computational basis. Thus the state of $(\mathcal{P}_C, \mathcal{A}_C)$ becomes

$$\begin{aligned} C_k &:= \mathcal{M}_{\bar{S}_k} \left(\frac{1}{\beta'} \Pi_{b_k, d_k} U_P \left(\mathcal{M}_{msg} (\Pi_{a'_k} U'_k C'_{k-0.5} U_k'^{\dagger} \Pi_{a'_k}^{\dagger}) \otimes |0, r_k\rangle \langle 0, r_k|_{AR} \right) U_P^{\dagger} \Pi_{b_k, d_k}^{\dagger} \right) \\ &= \mathcal{M}_{\bar{S}_k} \left(\frac{1}{\beta'} \Pi_{b_k, d_k} U_P \left(\mathcal{M}_{msg} (\Pi_{a'_k} U'_k \mathcal{M}'_{\bar{S}_k} (|Q_{k-0.5}\rangle \langle Q_{k-0.5}|) U_k'^{\dagger} \Pi_{a'_k}^{\dagger}) \otimes |0, r_k\rangle \langle 0, r_k|_{AR} \right) \right. \\ &\quad \left. U_P^{\dagger} \Pi_{b_k, d_k}^{\dagger} \right) \end{aligned}$$

where β' is the probability of outputting a'_k, b_k , and d_k . Similar to Fail-stop case, $\mathcal{M}_{\bar{S}_k}$ and $U_P \Pi_{b_k, d_k}$ commute by Lemma 1. $\mathcal{M}'_{\bar{S}_k}$ and $\Pi_{a'_k} U'_k$ also commute because they act on \bar{S}_k and S_k separately. Thus

$$\begin{aligned} C_k &= \frac{1}{\beta'} \Pi_{b_k, d_k} U_P \mathcal{M}_{\bar{S}_k} \mathcal{M}_{msg} \mathcal{M}'_{\bar{S}_k} \left(\Pi_{a'_k} U'_k |Q_{k-0.5}\rangle \langle Q_{k-0.5}| U_k'^{\dagger} \Pi_{a'_k}^{\dagger} \otimes |0, r_k\rangle \langle 0, r_k|_{AR} \right) U_P^{\dagger} \Pi_{b_k, d_k}^{\dagger} \\ &= \frac{1}{\beta'} \Pi_{b_k, d_k} U_P \left(\mathcal{M}_{\bar{S}_k} (\Pi_{a'_k} U'_k |Q_{k-0.5}\rangle \langle Q_{k-0.5}| U_k'^{\dagger} \Pi_{a'_k}^{\dagger}) \otimes |0, r_k\rangle \langle 0, r_k|_{AR} \right) U_P^{\dagger} \Pi_{b_k, d_k}^{\dagger} \\ &= \mathcal{M}_{\bar{S}_k} \left(\frac{1}{\beta'} \Pi_{b_k, d_k} U_P \left(\Pi_{a'_k} U'_k |Q_{k-0.5}\rangle \langle Q_{k-0.5}| U_k'^{\dagger} \Pi_{a'_k}^{\dagger} \otimes |0, r_k\rangle \langle 0, r_k|_{AR} \right) U_P^{\dagger} \Pi_{b_k, d_k}^{\dagger} \right) \\ &= \frac{\beta}{\beta'} \mathcal{M}_{\bar{S}_k} (|Q_k\rangle \langle Q_k|). \end{aligned}$$

where the second equality is because $\mathcal{M}_{\bar{S}_k}$ measures a larger space than $\mathcal{M}_{msg} \mathcal{M}'_{\bar{S}_k}$, thus $\mathcal{M}_{\bar{S}_k} \mathcal{M}_{msg} \mathcal{M}'_{\bar{S}_k} \equiv \mathcal{M}_{\bar{S}_k}$.

Finally, since C_k has trace 1, we have $\beta = \beta'$ and thus $C_k = \mathcal{M}_{\bar{S}_k} (|Q_k\rangle \langle Q_k|)$. The probability of \mathcal{E}_k occurring is $p\gamma\beta$ for both cases. \blacktriangleleft