Quantum query complexity and distributed computing
Röhrig, H.P.

Citation for published version (APA):
Chapter 6

Quantum Coin Flipping

This chapter is based on joint research conducted with Ambainis, Buhrman, and Dodis [12].

6.1 Introduction

Research into quantum cryptography is motivated by two observations about quantum mechanics:

1. Nonorthogonal quantum states cannot be distinguished perfectly and parts of certain orthogonal quantum states cannot be distinguished if the remaining parts are inaccessible;

2. Measurement disturbs the quantum state. This is the so-called "collapse of the wave function."

The second observation hints at the possibility of detecting eavesdroppers or other types of cheaters, whereas the first property appears to allow hiding data. Both rely on assumptions about the physical world, but are unhampered by unproven computational assumptions. Indeed, for the task of cooperatively establishing a random bit string between two parties in the presence of eavesdroppers, quantum key distribution [21, 89, 84] achieves security against the most general attack by an adversary that has unbounded computational power but has to obey the laws of quantum mechanics.

Initially, it was thought that these properties would admit protocols for the cryptographic primitive bit commitment. In bit commitment, there are two parties Alice and Bob; in the initial phase of the protocol, Alice has a bit $b$ and communicates with Bob to "commit" to the value of $b$ without revealing it. At a later time, Alice "unveils" her bit, allowing Bob to perform checks against the information obtained in the initial phase to test whether
Chapter 6. Quantum coin flipping

the revealed bit equals the committed bit. The properties sought of bit-commitment protocols are that they are concealing and binding: Bob does not learn anything about $b$ in the initial phase and Bob will catch Alice trying to unveil $1 - b$ instead of $b$.

Unfortunately, Mayers [90] and Lo and Chau [83] proved that perfect quantum bit commitment is impossible. Their impossibility result extends to strong coin tossing [91, 83], a weaker cryptographic primitive where the two parties want to agree on a random bit whose value cannot be influenced by either of them. Moreover, the impossibility extends even to the case of weak coin tossing [10], where outcome $b = 0$ is favorable for Alice and outcome $b = 1$ favorable for Bob, thus ruling out perfect quantum protocols for leader election. However, what turned out to be possible are coin-tossing protocols where there are guarantees on how much a cheater can bias the outcome.

Consider $k$ parties out of which at least $g \geq 1$ are honest and at most $(k - g)$ are dishonest; which players are dishonest is fixed in advance but unknown to the honest players. The players can communicate over broadcast channels. Initially they do not share randomness, but they can privately flip coins; the probabilities below are with respect to the private random coins. A coin-flipping protocol establishes among the honest players a bit $b$ such that

- if all players are honest, $\Pr[b = 0] = \Pr[b = 1] = 1/2$
- if at least $g$ players are dishonest, then $\Pr[b = 0], \Pr[b = 1] \leq 1/2 + \varepsilon$

$\varepsilon$ is called the bias; a small bias implies that colluding dishonest players cannot strongly influence the outcome of the protocol. Players may abort the protocol. This allows the bad players to block outcomes they do not desire; therefore the quality of a coin-flipping protocol is measured in terms of the overall probability of forcing a fixed outcome. Frequent aborts reduce this figure of merit.

Classically, if a weak majority of the players is bad then no bias $< 1/2$ can be achieved and hence no meaningful protocols exist [104]. For example, if we only have two players and one of them is dishonest, then no protocols with bias $< 1/2$ exist. For a minority of bad players, quite non-trivial protocols exist. For example, Feige [52] elegantly showed that $(1/2 + \delta)$-fraction of good players can achieve bias $\frac{1}{2} - \Omega(\delta^{1.65})$, while achieving bias better than $\frac{1}{2} - \delta$ is impossible.

Allowing qubits to be sent instead of classical bits changes the situation dramatically. Surprisingly, already in the two-party case coin flipping with bias $< 1/2$ is possible, as was first shown in [4]. The best known bias is $1/4$ and this is optimal for a special class of three-round protocols [10]; for a bias of $\varepsilon$ at least $\Omega(\log \log (1/\varepsilon))$ rounds of communication are necessary
Kitaev (unpublished, see [79]) showed that in the two-party case no bias smaller than $1/\sqrt{2} - 1/2$ is possible.

A weak version of the coin-flipping problem is one in which we know in advance that outcome 0 benefits Alice and outcome 1 benefits Bob. In this case, we only need to bound the probabilities of a dishonest Alice convincing Bob that the outcome is 0 and a dishonest Bob convincing Alice that the outcome is 1. In the classical setting, a standard argument shows that even weak coin flipping with a bias $< 1/2$ is impossible when a majority of the players is dishonest. In the quantum setting, this scenario was first studied under the name quantum gambling [63]. Subsequently, Spekkens and Rudolph [111] gave a quantum protocol for weak coin flipping with bias $1/\sqrt{2} - 1/2$, i.e., no party can achieve the desired outcome with probability greater than $1/\sqrt{2}$. Notice that this is a better bias than in the best strong coin flipping protocol of [10].

We also remark that Kitaev’s lower bound proof for strong coin flipping does not apply to weak coin flipping. Thus, weak protocols with arbitrarily small $\varepsilon > 0$ may be possible. The only known lower bounds for weak coin flipping are that the protocol of [111] is optimal for a restricted class of protocols [11] and that a protocol must use at least $\Omega(\log \log(1/\varepsilon))$ rounds of communication to achieve bias $\varepsilon$. This was shown in [10] for strong coin flipping but the proof also applies to weak coin flipping.

In this chapter, we focus on quantum coin flipping for more than two players. However, for our multiparty quantum protocols we will first need a new two-party quantum protocol for coin flipping with penalty for cheating. In this problem, players can be heavily penalized for cheating, which will allow us to achieve lower cheating probability as a function of the penalty. This primitive and the quantum protocol for it are presented in Section 6.2; they may be of independent interest.

One way to classically model communication between more than two parties is by a primitive called broadcast. When a player sends a bit to the other players he broadcasts it to all the other players at once [18]. However, when we deal with qubits such a broadcast channel is not possible since it requires to clone or copy the qubit to be broadcast and cloning a qubit is not possible [117]. In Section 6.3 we develop a proper quantum version of the broadcast primitive, which generalizes the classical broadcast. Somewhat surprisingly, we show that our quantum broadcast channel is essentially as powerful as a combination of pairwise quantum channels and a classical broadcast channel. This could also be of independent interest.

Using this broadcast primitive we obtain our main result:

6.1.1. **Theorem.** For $k$ parties out of which $g$ are honest, the optimal achievable bias is $\left(\frac{1}{2} - \Theta(\frac{g}{k})\right)$. 
We prove Theorem 6.1.1 by giving an efficient protocol with bias \( \frac{1}{2} - \Omega(\frac{1}{\sqrt{n}}) \) in Section 6.4 and showing a lower bound of \( \left( \frac{1}{2} - O\left(\frac{1}{\sqrt{n}}\right)\right) \) in Section 6.5. Our protocol builds upon our two-party coin-flipping with penalties which we develop in Section 6.2, and the classical protocol of Feige [52] which allows to reduce the number of participants in the protocol without significantly changing the fraction of good players present. Our lower bound extends the lower bound of Kitaev [79].

### 6.2 Two-Party Coin Flipping with Penalty for Cheating

We consider the following model for coin flipping. We have two parties: Alice and Bob, among at least one is assumed to be honest. If no party is caught cheating, the winner gets 1 coin, the loser gets 0 coins. If honest Alice catches dishonest Bob, Bob loses \( v \) coins but Alice wins 0 coins. Similarly, if honest Bob catches dishonest Alice, she loses \( v \) coins but Bob wins 0 coins.

#### 6.2.1 Theorem

If Alice (Bob) is honest, the expected win by dishonest Bob (Alice) is at most \( \frac{1}{2} + \frac{1}{\sqrt{v}} \), for \( v \geq 4 \).

**Proof.** The protocol is as follows. Let \( \delta = \frac{2}{\sqrt{v}} \). Define \( |\psi_a\rangle = \sqrt{\delta}|a\rangle|a\rangle + \sqrt{1 - \delta}|2\rangle|2\rangle \).

1. Alice picks \( a \in \{0, 1\} \) uniformly at random, generates the state \( |\psi_a\rangle \) and sends the second register to Bob.

2. Bob stores this state in a quantum memory, picks \( b \in \{0, 1\} \) uniformly at random and sends \( b \) to Alice.

3. Alice then sends \( a \) and the first register to Bob and Bob verifies if the joint state of the two registers is \( |\psi_a\rangle \) by measuring it in a basis consisting of \( |\psi_a\rangle \) and everything orthogonal to it. If the test is passed, the result of coin flip is \( a \oplus b \), otherwise Bob catches Alice cheating.

Theorem 6.2.1 follows from the following two claims.

#### 6.2.2 Claim

Bob cannot win with probability more than \( \frac{1}{2} + \frac{1}{\sqrt{v}} \), thus his expected win is at most \( \frac{1}{2} + \frac{1}{\sqrt{v}} \).

**Proof.** Let \( \rho_a \) be the density matrix of the second register of \( |\psi_a\rangle \). Then, for the trace distance between \( \rho_0 \) and \( \rho_1 \) we have \( ||\rho_0 - \rho_1||_1 = 2\delta \).
Aharonov et al. [3] showed that the trace distance is a measure for the distinguishability of quantum states analogously to the total variation distance of probability distributions; in particular, the probability of Bob winning is at most $\frac{1}{2} + \frac{\|\alpha - \beta\|_1}{4} = \frac{1}{2} + \frac{\varepsilon}{2} = \frac{1}{2} + \frac{1}{\sqrt{c}}$.

6.2.3. CLAIM. Dishonest Alice's expected win is at most $\frac{1}{2} + \frac{1}{\sqrt{c}}$.

**Proof.** Without loss of generality, we can assume that Alice is trying to achieve $a \oplus b = 0$, which is equivalent to $a = b$. Since initially she has no information about the $b$ that Bob is going to send, the state she sends in the first round is independent of $b$. So she prepares some pure quantum state $|\psi\rangle$, of which a part is sent to Bob. We can assume that this state is of the form

$$|\psi\rangle = \alpha_0 |0\rangle|0\rangle + \alpha_1 |1\rangle|1\rangle + \alpha_2 |2\rangle|2\rangle$$

for some $\alpha_0, \alpha_1, \alpha_2 \geq 0$, because all that matters is the purification of the density matrix that Bob receives. Moreover, by symmetry we can assume that the amplitudes $\alpha_1$ and $\alpha_2$ have the same magnitude so

$$|\psi\rangle = \sqrt{\varepsilon}|0\rangle|0\rangle + \sqrt{\varepsilon}|1\rangle|1\rangle + \sqrt{1-2\varepsilon}|2\rangle|2\rangle$$

for some $\varepsilon \geq 0$. Since the state is symmetric with respect to switching $|0\rangle$ and $|1\rangle$, the maximum expected win that Alice can achieve is the same is she receives $b = 0$ from Bob and if she receives $b = 1$.

It suffices to consider the case when she receives $b = 0$. After receiving $b = 0$, Alice performs a measurement on her register. By $|\psi'_0\rangle$ we denote the projection of $|\psi\rangle$ to the subspace in which Alice answers $a = 0$. Hence, $|\psi\rangle = |\psi'_0\rangle + |\psi'_1\rangle$. By symmetry, we can assume that

$$|\psi'_0\rangle = \sqrt{\varepsilon_0}|0\rangle|0\rangle + \sqrt{\varepsilon_1}|1\rangle|1\rangle + \sqrt{1-x-\varepsilon}|2\rangle|2\rangle,$$

$$|\psi'_1\rangle = \sqrt{\varepsilon - \varepsilon_0}|0\rangle|0\rangle + \sqrt{\varepsilon - \varepsilon_1}|1\rangle|1\rangle + \sqrt{1-x-\varepsilon}|2\rangle|2\rangle$$

for some $\varepsilon_0, \varepsilon_1, x \geq 0$. The best strategy for Alice is just to send the first register to Bob unchanged. The probability with which Alice succeeds is $|\langle \psi'_0 | \psi_0 \rangle |^2$ for $a = 0$ and $|\langle \psi'_1 | \psi_1 \rangle |^2$ for $a = 1$. If $\varepsilon_1 > 0$, then changing $\varepsilon_1$ to 0 does not change $|\langle \psi'_0 | \psi_0 \rangle |^2$ and increases $|\langle \psi'_1 | \psi_1 \rangle |^2$. Similarly, changing $\varepsilon_0$ to $\varepsilon$ does not change $|\langle \psi'_1 | \psi_1 \rangle |^2$ and increases $|\langle \psi'_0 | \psi_0 \rangle |^2$. Therefore, we can assume that $\varepsilon_0 = \varepsilon, \varepsilon_1 = 0$ and the states are

$$|\psi'_0\rangle = \sqrt{\varepsilon}|0\rangle|0\rangle + \sqrt{x-\varepsilon}|2\rangle|2\rangle,$$

$$|\psi'_1\rangle = \sqrt{\varepsilon}|1\rangle|1\rangle + \sqrt{1-x-\varepsilon}|2\rangle|2\rangle.$$
Let $|\psi_i''\rangle = \sqrt{1 - \delta}|i\rangle - \sqrt{\delta}|2\rangle$ for $i \in \{0, 1\}$. Then $|\psi_i''\rangle$ is orthogonal to $|\psi_i\rangle$, so we can assume that Bob's verification measurement has $|\psi_i''\rangle$ as one of the outcomes that indicate that Alice is cheating. Therefore, the probability of Alice caught cheating is at least $\langle\psi_0''|\psi_0''\rangle^2 + \langle\psi_1''|\psi_1''\rangle^2$.

Let $d = \max\{x, 1 - x\} - \frac{1}{2}$. Then the probability of Alice claiming $a = 0$ (and hence forcing outcome $a \oplus b = 0$ as desired) is $\langle\psi_0'|\psi_0'\rangle = x \leq \frac{1}{2} + d$. However, she may be caught cheating. We claim

6.2.4. Claim. The probability of Alice being caught by Bob is at least $\frac{d^2\delta}{2}$.

Proof. Consider the two inner products

$\langle\psi_0''|\psi_0'\rangle = \sqrt{\varepsilon}\sqrt{1 - \delta} - \sqrt{x - \varepsilon}\sqrt{\delta}$,

$\langle\psi_1''|\psi_1'\rangle = \sqrt{\varepsilon}\sqrt{1 - \delta} - \sqrt{1 - x - \varepsilon}\sqrt{\delta}$

To compare their difference, note that

$\sqrt{x - \varepsilon} - \sqrt{1 - x - \varepsilon} \geq \frac{x - (1 - x)}{\sqrt{x} + \sqrt{1 - x}} \geq \frac{x - (1 - x)}{\sqrt{2}}$

where the first inequality follows from convexity of square root function and the second inequality follows from Cauchy-Schwartz. Therefore, $\langle\psi_0''|\psi_0'\rangle$ and $\langle\psi_1''|\psi_1'\rangle$ differ in absolute value by at least $\frac{|x - (1 - x)|\sqrt{2}}{\sqrt{2}} = d\sqrt{2}\delta$. This implies that one of $|\langle\psi_0''|\psi_0'\rangle|^2$ and $|\langle\psi_1''|\psi_1'\rangle|^2$ is at least $\frac{d^2\delta}{2}$ and Alice gets caught with probability at least $\frac{d^2\delta}{2}$.

Therefore, Alice's expected win is at most

$\frac{1}{2} + d - \frac{d^2\delta v}{2} = \frac{1}{2} + d\left(1 - \frac{d\delta v}{2}\right)$.

Consider two cases. If $d\delta v \geq 2$, then $1 - \frac{d\delta v}{2} \leq 0$ and the expected win is at most $\frac{1}{2}$. If $d\delta v \leq 2$, then $d \leq \frac{2}{\delta v} = \frac{1}{\sqrt{v}}$ and Alice's expected win is at most $\frac{1}{2} + \frac{1}{\sqrt{v}}$. 

6.3 The Multiparty Model

6.3.1 Adversaries

We assume computationally unbounded adversaries. However, they have to obey quantum mechanics and cannot read the private memory of the honest.
players, but they can communicate secretly with each other. Moreover, we assume that they can only access the message space in between rounds or when according to the protocol it is their turn to send a message.

6.3.2 The broadcast channel

A classical broadcast channel allows one party to send a classical bit to all the other players. In the quantum setting this would mean that a qubit would be sent to all the other players. However, when there are more than two players in total we would have to \textit{clone} or \textit{copy} the qubit in order to send it to the other players. Even if the sender knows a classical preparation of the state he wants to send, we cannot allow him to prepare copies because he may be a cheater and send different states to different parties. It is well known that it is impossible to clone a qubit [117], because cloning is not a unitary operation. This means that we will have to take a slightly different approach. Quantum broadcast channels have been studied in an information-theoretic context before [14, 116] but not in the presence of faulty or malicious parties.

Our quantum broadcast channel works as follows. Suppose there are \( k \) players in total and that one player wants to broadcast a qubit that is in the state \( \alpha|0\rangle + \beta|1\rangle \). What will happen is that the channel will create the \( k \)-qubit state \( \alpha|0^k\rangle + \beta|1^k\rangle \) and send one of the \( k \) qubits to each of the other players. The state \( \alpha|0^k\rangle + \beta|1^k\rangle \) can be easily created from \( \alpha|0\rangle + \beta|1\rangle \) by taking \( k - 1 \) fresh qubits in the state \( |0^{k-1}\rangle \). This joint state can be written as \( \alpha|0^k\rangle + \beta|0^{k-1}\rangle \). Next we flip the last \( k - 1 \) bits conditional on the first bit being a 1, thus obtaining the desired state \( \alpha|0^k\rangle + \beta|1^k\rangle \). This last operation can be implemented with a series of controlled-not operations. Note that this state is not producing \( k \) copies of the original state, which would be the \( k \)-fold product state \((\alpha|0\rangle + \beta|1\rangle) \otimes \ldots \otimes (\alpha|0\rangle + \beta|1\rangle) \).

6.3.1 Theorem. In the following sense, a quantum broadcast channel between \( k \) parties is comparable to models where the parties have a classical broadcast channel and/or pairwise quantum channels:

- If all parties are honest:

1. One use of the quantum broadcast channel can be simulated with \( 2(k - 1) \) uses of pairwise quantum channels.

2. One use of a classical broadcast channel can be simulated with one use of the quantum broadcast channel.

3. One use of a pairwise quantum channel can be simulated by \( k + 1 \) uses of the quantum broadcast channel.
• If all but one of the parties are dishonest, using one of the simulations above in place of the original communication primitive does not confer extra cheating power.

Proof. We first give the simulations and argue that they work in case all players are honest.

1. The sender takes \( k - 1 \) fresh qubits in state \(|0^k\rangle\). He applies \( k - 1 \) times CNOT where the subsystem to be broadcast is the control of the CNOT and the fresh qubits are the destination. He then sends each of the \( k - 1 \) qubits via the pairwise quantum channels to the \( k - 1 \) other parties. Each recipient \( j \) flips a private classical random bit \( r_j \) and if \( r_j = 1 \) performs a \( \sigma_z \) phase flip on the received qubit. Here \( \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \) is the Pauli matrix that multiplies the relative phase between the \(|0\rangle\) and the \(|1\rangle\) state by \(-1\). He then sends \( r_j \) back to the sender. The sender computes the parity of the \( r_j \) and if it is odd, he performs a \( \sigma_z \) phase flip on his part of the broadcast state, thus restoring the correct relative phase. This randomization is a countermeasure; its utility is explained below.

2. When the sender wants to broadcast bit \( b \in \{0, 1\} \), he uses the quantum broadcast channel on qubit \(|b\rangle\). The recipients immediately measure their qubit in the computational basis to obtain the classical bit.

3. The quantum broadcast channel can be used to create an EPR pair \((|00\rangle + |11\rangle)/\sqrt{2}\) between two players \(P_i\) and \(P_j\) with the assistance of the other \((k - 2)\) players. \(i\) and \(j\) are determined by the protocol.

First one player broadcasts the state \((|0\rangle + |1\rangle)/\sqrt{2}\), resulting in the \(k\) qubit state \(|\varphi\rangle = (|0^k\rangle + |1^k\rangle)/\sqrt{2}\). Now one after the other, the \(k - 2\) remaining players perform a Hadamard transformation on their qubit, measure it in the computational basis, and broadcast the classical result. Next, if \(P_i\) receives a 1 he applies a phase flip \(\sigma_z\) to his part of \(|\varphi\rangle\) (\(P_j\) does nothing). After this operation, \(|\varphi\rangle\) will be an EPR state between \(P_i\) and \(P_j\) unentangled with the other \(k - 2\) parties. Using a shared EPR pair, a protocol called teleportation [19] can be used to simulate a private quantum channel between \(P_i\) and \(P_j\). Teleportation requires the transmission of two bits of classical information.

For the case of all but one party being dishonest:

1. If the sender is honest, the recipients obtain exactly the same subsystems as for the quantum broadcast channel.
If one of the recipients is honest, he may receive an arbitrary quantum subsystem up to the randomized relative phase. However, exactly the same can be achieved with a quantum broadcast channel with $k - 1$ cheating parties, who each perform a Hadamard transformation on their subsystem followed by a measurement in the computational basis.

2. If the sender is honest, all recipients obtain the same computational-basis state.

If one of the recipients is honest, he obtains a classical bit that is possibly randomized in case the dishonest sender does not broadcast a basis state. Since the sender can flip a coin himself, this does not give more cheating power.

3. If the sender is honest, we can assume without loss of generality that all cheating action is done after the EPR pair has been established, because the merged cheaters can easily recreate the original broadcast state and also compensate phase flips of the honest sender. However, after the EPR pair has been established, the sender unilaterally performs his part of the teleportation circuit and measurements and sends the two bits of classical information. So the most general cheating action is to apply a quantum operation after the reception of the two classical bits. Furthermore, we can even assume that the cheating action is done after the correction circuit of teleportation. This is similar to the teleportation of quantum gates [67], and, hence, amounts to cheating on a pairwise quantum channel.

If one of the recipients is honest, the best the cheaters can aim for is to give an arbitrary quantum state to the honest recipient. This they can also achieve over a pairwise quantum channel.

\[ \square \]

### 6.4 Multiparty Quantum Coin-Flipping Protocols

We will first consider the case of only one good player, i.e., $g = 1$, and later extend our results to general $g$.

**One honest player** Recall, we need to construct a protocol with bias $1/2 - \Omega(1/k)$. Before proceeding to our actual protocol, let us consider a simple protocol which trivially extends the previous work in the two-party setting,
but does not give us the desired result. The protocols is as follows: player 1 flips a random coin with player 2, player 3 flips a random coin with player 4 and so forth. In each pair, the player with the higher id wins if the coin is 1 and the one with the lower id if the coin is 0. The winners repeat the procedure. With each repetition of the tournament, half of the remaining players are eliminated. If there is an odd number of players at any moment, the one with the highest id advances to the next round. When there are only two players left, the coin they flip becomes the output of the protocol. Above we assume we have private point-to-point quantum channels and a classical broadcast channel, which is justified by Theorem 6.3.1.

Now, the elimination rounds can be implemented using the weak two-party coin-tossing protocol by Spekkens and Rudolph [111] and the last round by the the strong two-party coin-tossing protocol by Ambainis [10]. If there is only one good player, the probability that he makes it to the last round is \((1 - 1/\sqrt{2})^{-1+\log k}\); in this case, the probability that the bad players can determine the output coin is \(3/4\). In case the good player gets eliminated, the bad players can completely determine the coin. Hence, the overall probability that the bad players can determine the coin is \(1 - \frac{1}{4}(1 - \frac{1}{\sqrt{2}})^{-1+\log k} \leq 1 - \frac{1}{4^{k^{1.78}}},\) which corresponds to bias \(\frac{1}{2} - \Omega(1/k^{1.78})\).

To improve the above naive bound to the desired value \(\frac{1}{2} - \Omega(1/k)\), we will use our coin-flipping protocol with penalty from Section 6.2. The idea is that in current quantum coin-flipping protocols for two parties, there are three outcomes for a given player: "win," "lose," and "abort." Now, looking at the elimination tournament above, if an honest player loses a given coin flipping round, he does not "complain" and bad player win the game. However, if the honest player detect cheating, he can and will abort the entire process, which corresponds to the failure of the dishonest players to fix the coin. Of course, if the are few elimination rounds left, bad players might be willing to risk the abort if they gain significant benefits in winning the round. However, if the round number is low, abort becomes prohibitively expensive: a dishonest player might not be willing to risk it given there are plenty more opportunities for the honest player to "normally lose". Thus, instead of regular two-party coin-tossing protocols, which do not differentiate between losing and abortion, we can employ our protocol for coin flipping with penalty, where the penalties are very high at the original rounds, and eventually get lower towards the end of the protocol. Specific penalties are chosen in a way which optimizes the final bias we get, and allows us to achieve the desired bias \(1/2 - \Omega(1/k)\).

6.4.1. THEOREM. *There is a strong quantum coin-tossing protocol for k parties with bias at most \(1/2 - c/k\) for some constant c, even with \((k - 1)\) bad parties.*

**Proof.** We assume that \(k = 2^n\) for some \(n > 0\), as it changes c by at most
6.4. Multiparty quantum coin-flipping protocols

a constant factor. Let $Q_v$ be the maximum expected win in a two-party protocol with penalty $v$. Consider the following protocol with $n$ rounds.

In the $i$th round, we have $2^{n+1-i}$ parties remaining. We divide them into pairs. Each pair performs the two-party coin-flipping protocol with penalty $(2^{n-i} - 1)$, with Alice winning if the outcome is 1 and Bob winning if the outcome is 0. The winners proceed to $(i+1)$th round.

In the $(n-2)$nd round, there are just 8 parties remaining. At this stage, they can perform three rounds of regular coin flipping with no penalty of [10, 77] in which no cheater can bias the coin to probability more than $3/4$, which will result in maximum probability of $63/64$ of fixing the outcome. The result of this last two-round protocol is the result of our $2^n$-party protocol.

Assume that the honest player has won the first $(n-j)$ coin flips and advanced to $(j+1)$th round. Assume that the all other players in the $(j+1)$th round are dishonest. Let $P_j$ be the maximum probability with which $(2^j - 1)$ dishonest players can fix the outcome to 0 (or 1).

6.4.2. Claim.

$$1 - P_j \geq (1 - P_{j-1})(1 - Q_{2^{j-1}-1})$$ (6.1)

Proof. Let $p_w$, $p_l$, $p_c$ be the probabilities of the honest player winning, losing and catching the other party cheating in the $(j+1)$th round of the protocol. Notice that $p_w + p_l + p_c = 1$. Then, the probability $P_j$ of $2^j - 1$ dishonest parties fixing the coin is at most $p_l + p_w P_{j-1}$. If the honest player loses, they win immediately. If he wins, they can still bias the coin in $j-1$ remaining rounds to probability at most $P_{j-1}$. If he catches his opponent cheating, he exits the protocol and the dishonest players have no more chances to cheat him. Using $p_w = 1 - p_l - p_c$, we have

$$P_j \leq p_l + p_w P_{j-1} = P_{j-1} + (1 - P_{j-1})p_l - P_{j-1}p_c$$

$$= P_{j-1} + (1 - P_{j-1}) \left( p_l - \frac{P_{j-1}}{1 - P_{j-1}} p_c \right)$$ (6.2)

Next, notice that $P_{j-1} \geq 1 - \frac{1}{2^{j-1}}$. This is because $2^{j-1} - 1$ bad players could just play honestly when they face the good player and fix the coin flip if two bad players meet in the last round. Then, the probability of the good player winning all $j-1$ rounds is $\frac{1}{2^{j-1}}$. Therefore, $\frac{P_{j-1}}{1 - P_{j-1}} \geq 2^{j-1} - 1$ and (6.2) becomes

$$P_j \leq P_{j-1} + (1 - P_{j-1})(p_l - (2^{j-1} - 1)p_c)$$ (6.3)

Finally, the term in brackets is at most $Q_{2^{j-1}-1}$, which gives

$$P_j \leq P_{j-1} + (1 - P_{j-1})Q_{2^{j-1}-1}$$ (6.4)

which in turn is equivalent to the desired (6.1).
By applying the claim inductively, we get

\[ 1 - P_n \geq \frac{1}{64} \prod_{j=4}^{n} (1 - Q_{2^j-1-1}) \]

where the \( \frac{1}{64} \) term comes from the naive protocol we use in the last three rounds. Now, using the bound in Theorem 6.2.1 we have

\[
1 - P_n \geq \frac{1}{64} \prod_{j=3}^{n-1} (1 - Q_{2^j-1}) \geq \frac{1}{64} \prod_{j=3}^{n-1} \left( \frac{1}{2} - \frac{1}{\sqrt{2^j - 1}} \right) \\
\geq \frac{1}{8 \cdot 2^n} \prod_{j=3}^{n-1} \left( 1 - \frac{2}{\sqrt{2^j - 1}} \right).
\]

The last term in the brackets is at least \( \prod_{j=3}^{\infty} (1 - \frac{2}{\sqrt{2^j - 1}}) \) which is a positive constant. Therefore, for some constant \( c > 0 \) we have \( 1 - P_n \geq \frac{c}{2^n} = \frac{c}{k} \), which means that the bias is at most \( \frac{1}{2} - \Omega(\frac{1}{k}) \).

**Extending to many honest players** We can extend Theorem 6.4.1 to every number \( g \geq 1 \) of good players by using the classical lightest-bin protocol of Feige [52]. This protocol allows us to reduce the total number of players until a single good player is left without significantly changing the fraction of good players, after which we can run the quantum protocol of Theorem 6.4.1 to get the desired result. Specifically, Lemma 8 from [52] implies that starting from \( g = \delta k \) good players out of \( k \) players, the players can classically select a sub-committee of \( O(1/\delta) = O(k/g) \) players containing at least one good player with probability at least 1/2. Now, this sub-committee can use the quantum protocol of Theorem 6.4.1 to flip a coin with bias \( 1/2 - \Omega(g/k) \), provided it indeed contains at least one honest player. But since the latter happens with probability at least 1/2, the final bias is at most \( 1/2 - (1/2) \cdot \Omega(g/k) = 1/2 - \Omega(g/k) \), as desired.

### 6.5 Lower Bound

#### 6.5.1 The two-party bound

For completeness and to facilitate the presentation of our generalization, we reproduce here Kitaev's unpublished proof [79] that every two-party strong quantum coin-flipping protocol must have bias at least \( 1/\sqrt{2} \). The model here is that the two parties communicate over a quantum channel.
6.5. Lower bound

6.5.1. Definition. Let \( \mathcal{H} := \mathcal{A} \otimes \mathcal{M} \otimes \mathcal{B} \) denote the Hilbert space of the coin-flipping protocol composed of Alice’s private space, the message space, and Bob’s private space. A 2N-round two-party coin-flipping protocol is a tuple

\[
(U_{A,1}, \ldots, U_{A,N}, U_{B,1}, \ldots, U_{B,N}, \Pi_{A,0}, \Pi_{A,1}, \Pi_{B,0}, \Pi_{B,1})
\]

where

- \( U_{A,j} \) is a unitary operator on \( \mathcal{A} \otimes \mathcal{M} \) for \( j = 1, \ldots, N \),
- \( U_{B,j} \) is a unitary operator on \( \mathcal{M} \otimes \mathcal{B} \) for \( j = 1, \ldots, N \),
- \( \Pi_{A,0} \) and \( \Pi_{A,1} \) are projections from \( \mathcal{A} \) onto orthogonal subspaces of \( \mathcal{A} \), representing Alice’s final measurements for outcome 0 and 1, respectively,
- \( \Pi_{B,0} \) and \( \Pi_{B,1} \) are projections from \( \mathcal{B} \) onto orthogonal subspaces of \( \mathcal{B} \), representing Bob’s final measurements for outcome 0 and 1, respectively,

so that for

\[
|\psi_N\rangle := (1_A \otimes U_{B,N})(U_{A,N} \otimes 1_B)(1_A \otimes U_{B,N-1})(U_{A,N-1} \otimes 1_B) \cdots \\
\cdots (1_A \otimes U_{B,1})(U_{A,1} \otimes 1_B)|0\rangle
\]

holds

\[
\langle \Pi_{A,0} \otimes 1_M \otimes 1_B | \psi_N \rangle = (1_A \otimes 1_M \otimes \Pi_{B,0})|\psi_N\rangle \tag{6.5}
\]

\[
\langle \Pi_{A,1} \otimes 1_M \otimes 1_B | \psi_N \rangle = (1_A \otimes 1_M \otimes \Pi_{B,1})|\psi_N\rangle \tag{6.6}
\]

\[
\|\langle \Pi_{A,0} \otimes 1_M \otimes 1_B | \psi_N \rangle\| = \|\langle \Pi_{A,1} \otimes 1_M \otimes 1_B | \psi_N \rangle\| \tag{6.7}
\]

The first two conditions ensure that when Alice and Bob are honest, they both get the same value for the coin and the third condition guarantees that when Alice and Bob are honest, their coin is not biased. A player aborts if her or his final measurement does not produce outcome 0 or 1; of course, it is no restriction to delay this action to the end of the protocol.

6.5.2. Lemma. Fix an arbitrary two-party quantum coin-flipping protocol. Let \( p_{1*} \) and \( p_{*1} \) denote the probability that Alice or Bob, respectively, can force the outcome of the protocol to be 1 if the other party follows the protocol. Denote by \( p_1 \) the probability for outcome 1 when there are no cheaters. Then \( p_{1*}p_{*1} \geq p_1 \).

Hence, if \( p_1 = 1/2 \), then \( \max\{p_{1*}, p_{*1}\} \geq 1/\sqrt{2} \). To prove Lemma 6.5.2, we construct the view of a run of the protocol from an honest Alice’s point
Chapter 6. Quantum coin flipping

of view, with Bob wanting to bias the protocol towards 1. The problem of optimizing Bob's strategy is a semidefinite program (SDP).

Semidefinite programming is a generalization of linear programming. In addition to the usual linear constraints, it is allowed to require that a square matrix of variables is positive semidefinite, i.e., all its eigenvalues are non-negative. The proof below makes use of the well-developed duality theory for SDPs. Let $A$, $B$, and $C$ denote square matrices of the same dimension. If $A$ is positive semidefinite, we write $A \geq 0$. We define $A \geq B :\iff A - B \geq 0$. The following properties are straightforward to verify:

$$ A \geq B \iff \forall \langle \psi \rangle : \langle \psi | A | \psi \rangle \geq \langle \psi | B | \psi \rangle $$
$$ A \geq B \iff \text{tr}_\mathcal{V} A \geq \text{tr}_\mathcal{V} B \text{ for every subspace } \mathcal{V} $$
$$ A = B + C \text{ and } C \geq 0 \iff A \geq B $$

6.5.3. Lemma. The optimal strategy of Bob trying to force outcome 1 is the solution to the following SDP over the semidefinite matrices $\rho_{A,0}, \ldots, \rho_{A,N}$ operating on $A \otimes \mathcal{M}$:

maximize \[ \text{tr} \left( (\Pi_{A,1} \otimes \mathds{1}_\mathcal{M}) \rho_{A,N} \right) \]
subject to

\[ \text{tr}_\mathcal{M} \rho_{A,0} = \langle 0 | \langle 0 |_A \]
\[ \text{tr}_\mathcal{M} \rho_{A,j} = \text{tr}_\mathcal{M} U_{A,j} \rho_{A,j-1} U_{A,j}^* \quad (1 \leq j \leq N) \]

Proof. Alice starts with her private memory in state $\langle 0 \rangle_A$ and we permit Bob to determine the $\mathcal{M}$ part of the initial state. Therefore all Alice knows is that initially, the space accessible to her is in state $\rho_{A,0}$ with $\text{tr}_\mathcal{M} \rho_{A,0} = \langle 0 | \langle 0 |_A$. Alice sends the first message, transforming the state to $\rho'_{A,0} := U_{A,1} \rho_{A,0} U_{A,1}^*$. Now Bob can do an arbitrary unitary operation on $\mathcal{M} \otimes \mathcal{B}$ leading to $\rho_{A,1}$, so the only constraint is $\text{tr}_\mathcal{M} \rho_{A,1} = \text{tr}_\mathcal{M} \rho'_{A,0}$. In the next round, honest Alice applies $U_{A,2}$, then Bob can do some operation that preserves the partial trace, and so forth. The probability for Alice outputting 1 is $\text{tr}((\Pi_{A,1} \otimes \mathds{1}_\mathcal{M}) \rho_{A,N})$ because the final state for Alice is $\rho_{A,N}$ and she performs an orthogonal measurement on $A$ with projections $\Pi_{A,0}, \Pi_{A,1}$, and $1_A - \Pi_{A,0} - \Pi_{A,1}$, which represents "abort."

6.5.4. Lemma. The dual SDP to the primal SDP in Lemma 6.5.3 is

minimize \[ \langle 0 | Z_{A,0} | 0 \rangle \]
subject to

\[ Z_{A,j} \otimes \mathds{1}_\mathcal{M} \geq U_{A,j+1}^* (Z_{A,j+1} \otimes \mathds{1}_\mathcal{M}) U_{A,j+1} \quad (0 \leq j \leq N - 1) \]
\[ Z_{A,N} = \Pi_{A,1} \]

over the Hermitian matrices $Z_{A,0}, \ldots, Z_{A,N}$ operating on $A$. 

\[ (6.8) \]
\[ (6.9) \]
\[ (6.10) \]
6.5. Lower bound

Proof. In the Lagrange-multiplier approach, a "primal" optimization problem

$$\max_{x \geq 0} f(x) \text{ subject to } g(x) \leq a \text{ with } a > 0$$

reformulated as

$$\max_{x} \inf_{\lambda \geq 0} f(x) - \lambda \cdot (g(x) - a),$$

which is bounded from above by \(\min_{\lambda \geq 0} \lambda \cdot a\) subject to \((f - \lambda \cdot g)(x) \leq 0\) for all \(x \geq 0\). In linear programming, \((f - \lambda \cdot g)(x) \leq 0\) for all \(x \geq 0\) if and only if \(f - \lambda \cdot g \leq 0\), therefore the preceding optimization problem can be simplified to \(\min_{\lambda \geq 0} \lambda \cdot a\) subject to \(f - \lambda \cdot g \leq 0\). The same construction can be applied to SDPs; we form the dual of the SDP in Lemma 6.5.3 as follows: the dual is equivalent to maximizing over the \(\rho_{A,j}\) the minimum of

$$\text{tr}((\Pi_{A,1} \otimes \mathcal{I})\rho_{A,N}) - \text{tr}(Z_{A,0}(\text{tr}_\mathcal{M} \rho_{A,0} - |0\rangle\langle 0|_A))$$

$$- \sum_{j=1}^{N} \text{tr}(Z_{A,j} (\text{tr}_\mathcal{M} (\rho_{A,j} - U_{A,j} \rho_{A,j-1} U_{A,j}^*)) - \sum_{j=0}^{N} \text{tr}(Y_j \rho_{A,j}) \quad (6.14)$$

subject to the operators \(Z_{A,j}\) on \(\mathcal{M}\) being Hermitian and the operators \(Y_j\) on \(\mathcal{A} \otimes \mathcal{M}\) being positive semidefinite, for \(0 \leq j \leq N\). In the above sum, the terms containing \(\rho_{A,j}\) for \(0 \leq j < N\) are

$$- \text{tr}(Z_{A,j} (\text{tr}_\mathcal{M} \rho_{A,j})) + \text{tr}(Z_{A,j+1} (\text{tr}_\mathcal{M} (U_{A,j+1} \rho_{A,j} U_{A,j+1}^*))) - \text{tr}(Y_j \rho_{A,j}) =$$

$$\text{tr}((-Z_{A,j} \otimes \mathcal{I}) + U_{A,j+1}^* (Z_{A,j+1} \otimes \mathcal{I}) U_{A,j+1} - Y_j) \rho_{A,j}\quad (6.15)$$

Since the primal constraints (6.9) and (6.10) are equality constraints, the dual constraint (6.15) must be equal to 0. However, since \(Y_j\) is positive semidefinite and does not appear anywhere else, we can drop it from (6.15) to arrive at the inequality (6.12).

For \(j = N\), we obtain the dual equality constraint (6.13) and the dual objective function becomes the only summand of (6.14) that does not involve any \(\rho_{A,j}\). \(\Box\)

Proof of Lemma 6.5.2. Let \(Z_{A,j}\) and \(Z_{B,j}\) \((0 \leq j \leq N)\) denote the optimal solutions for the dual SDPs for a cheating Bob and a cheating Alice, respectively. For each \(j\), \(0 \leq j \leq N\), let \(|\psi_j\rangle := (1_\mathcal{A} \otimes U_{B,j}) (U_{A,j} \otimes 1_B) \cdots (1_\mathcal{A} \otimes U_{B,1}) (U_{A,1} \otimes 1_B) |0\rangle\) denote the state of the protocol in round \(j\) when both parties are honest. Let \(F_j := \langle \psi_j | (Z_{A,j} \otimes \mathcal{I} \otimes Z_{B,j}) | \psi_j \rangle\). We claim

$$p_{1+} p_{+1} = F_0 \quad (6.16)$$

$$F_j \geq F_{j+1} \quad (0 \leq j < N) \quad (6.17)$$

$$F_N = p_1. \quad (6.18)$$
Combining (6.16)–(6.18), we obtain the desired $p_{1*}p_{*1} \geq p_1$. We now proceed to prove these claims.

Note that the primal SDP from Lemma 6.5.3 is strictly feasible: Bob playing honestly yields a feasible solution that is strictly positive. The strong-duality theorem of semidefinite programming states that in this case, the optimal value of the primal and the dual SDPs are the same, and therefore $p_{1*} = \langle 0|_A Z_{A,0}|0\rangle_A$ and $p_{*1} = \langle 0|_B Z_{B,0}|0\rangle_B$ and

$$p_{1*}p_{*1} = \langle 0|_A Z_{A,0}|0\rangle_A \cdot \langle 0|_M 1_M|0\rangle_M \cdot \langle 0|_B Z_{B,0}|0\rangle_B$$

$$= \langle 0|(Z_{A,0} \otimes 1_M \otimes Z_{B,0})|0\rangle = F_0.$$ 

The inequalities (6.17) hold because of the constraints (6.12). Equality (6.18) holds by constraint (6.13) we have

$$\langle \varphi|(Z_{A,N} \otimes 1_M \otimes Z_{B,N})|\varphi\rangle = \|((\Pi_{A,1} \otimes 1_M \otimes 1_B)(1_A \otimes 1_M \otimes \Pi_{B,1})|\varphi\rangle\|^2$$

for every $|\varphi\rangle$; $|\psi_N\rangle$ is the final state of the protocol when both players are honest, so by equation (6.6),

$$\|((\Pi_{A,1} \otimes 1_M \otimes 1_B)(1_A \otimes 1_M \otimes \Pi_{B,1})|\psi_N\rangle\|^2 = \|((\Pi_{A,1} \otimes 1_M \otimes 1_B)|\psi_N\rangle\|^2 = p_1.$$ 

\[ \square \]

### 6.5.2 More than two parties

We will now extend Kitaev's lower bound to $k$ parties. As with the upper bounds, we first start with a single honest player ($g = 1$), and then extend the result further to every $g$.

**6.5.5. Theorem.** Every strong quantum coin-tossing protocol for $k$ parties has bias at least $1/2 - (\ln 2)/k - O(1/k^2)$ if it has to deal with up to $(k - 1)$ bad parties.

We consider the model of private pairwise quantum channels between the parties; by Theorem 6.3.1 the results immediately carry over to the quantum broadcast channel. Before proving Theorem 6.5.5, we make the following detour.

**6.5.6. Definition.** Let $\mathcal{H} := A_1 \otimes \cdots \otimes A_k \otimes \mathcal{M}$ denote the Hilbert space composed of the private spaces of $k$ parties and the message space. An $N$-round $k$-party coin-flipping protocol is a tuple

$$(i_1, \ldots, i_N, U_1, \ldots, U_N, \Pi_{1,0}, \Pi_{1,1}, \ldots, \Pi_{k,0}, \Pi_{k,1})$$

where
6.5. **Lower bound**

- $i_j$ with $1 \leq i_j \leq k$, $1 \leq j \leq N$, indicates whose turn it is to access the message space in round $j$,
- $U_j$ is a unitary operator on $A_{i_j} \otimes M$ for $j = 1, \ldots, N$,
- for $1 \leq i \leq k$, $\Pi_{i,0}$ and $\Pi_{i,1}$ are projections from $A_i$ to orthogonal subspaces of $A_i$, representing the measurement that party $i$ performs to determine outcome 0 or 1, respectively,

so that for $|\psi_N\rangle := \tilde{U}_{i_N} \cdots \tilde{U}_{i_1} |0\rangle$ and each pair $1 \leq i < i' \leq k$ and every $b \in \{0,1\}$ holds

\[
\tilde{\Pi}_{i,b} |\psi_N\rangle = \tilde{\Pi}_{i',b} |\psi_N\rangle \quad (6.19)
\]
\[
||\tilde{\Pi}_{i,b} |\psi_N\rangle|| = ||\tilde{\Pi}_{i,1-b} |\psi_N\rangle||. \quad (6.20)
\]

Here $\tilde{U}_j$ denotes the extension of $U_j$ to all of $\mathcal{H}$ that acts as identity on the tensor factors $A_i$ for $i' \neq i$; $\tilde{\Pi}_{i,b} := (1_{A_i} \otimes \cdots \otimes 1_{A_i} \otimes \Pi_{i,b} \otimes 1_{A_{i+1}} \otimes \cdots \otimes 1_{A_k})$ is the extension of $\Pi_{i,b}$ to $\mathcal{H}$.

**6.5.7. LEMMA.** Fix an arbitrary quantum coin flipping protocol. For $b \in \{0,1\}$, let $p_b$ be the probability of outcome $b$ in case all players are honest. Let $p_{i,b}$ denote the probability that party $i$ can be convinced by the other parties that the outcome of the protocol is $b \in \{0,1\}$. Then

\[
p_{1,b} \cdots p_{k,b} \geq p_b
\]

**Proof of Lemma 6.5.7.** The optimal strategy for $k-1$ bad players trying to force outcome 1 is the solution to the SDP from Lemma 6.5.3 where all the cheating players are merged into a single cheating player.

Let $(Z_{i,j})_{0 \leq j \leq N}$ denote the optimal solution for the dual SDP for good player $i$, $1 \leq i \leq k$. For each $j$, $0 \leq j \leq N$, let $|\psi_j\rangle := \tilde{U}_j \cdots \tilde{U}_1 |0\rangle$ denote the state of the protocol in round $j$ when all parties are honest. Let $F_j := \langle \psi_j | (Z_{1,j} \otimes \cdots \otimes Z_{k,j} \otimes 1_M) |\psi_j\rangle$. By a similar argument as in the proof of Lemma 6.5.2, we have

\[
p_{1,1} \cdots p_{k,1} = F_0 \quad (6.21)
\]
\[
F_j \geq F_{j+1} \quad (0 \leq j < N) \quad (6.22)
\]
\[
F_N = p_1 \quad (6.23)
\]

Hence, $p_{1,1} \cdots p_{k,1} \geq p_1$. Repeating the argument with the cheaters aiming for outcome 0 completes the proof.

Now, Theorem 6.5.5 is an immediate consequence.
Proof of Theorem 6.5.5. Using the notation of Lemma 6.5.7, we have $p_0 = 1/2$. Let $q = \max_i p_{i,0}$ denote the maximum probability of any player forcing output 0. By Lemma 6.5.7, $q^k \geq p_{1,0} \cdots p_{k,0} \geq 1/2$, from which follows that $q \geq (1/2)^{1/k} \geq 1 - (\ln 2)/k - O(1/k^2)$. By Theorem 6.3.1 this result applies both to private pairwise quantum channels and the quantum broadcast channel.

Extending to many honest players Extension to any number of honest players follows almost immediately from Theorem 6.5.5. Indeed, take a protocol $\Pi$ for $k$ parties tolerating $(k - g)$ cheaters. Arbitrarily partition our players into $k' = k/g$ groups and view each each as one “combined player.” We get an induced protocol $\Pi'$ with $k'$ “super-players” which achieves at least the same bias $\varepsilon$ as $\Pi$, and can tolerate up to $(k' - 1)$ bad players. By Theorem 6.5.5, $\varepsilon \geq 1/2 - O(1/k') = 1/2 - O(g/k)$.

6.6 Summary

We showed that quantum coin flipping is significantly more powerful than classical coin flipping. Moreover, we give tight tradeoffs between the number of cheaters tolerated and the bias of the resulting coin achievable by quantum coin-flipping protocols. We also remark that the fact that we obtain tight bounds in the quantum setting is somewhat surprising. For comparison, such tight bounds are unknown for the classical setting.