On Bit-Commitment based Quantum Coin Flipping

Ashwin Nayak [*] — Caltech

Peter Shor [†] — AT&T Labs

Prepared October, 2001

Submitted April, 2002

Technical Report caltechCSTR/2002.004

Computer Science Department

California Institute of Technology

[*] Computer Science Department, and Institute for Quantum Information, California Institute of Technology, Mail Code 256-80, Pasadena, CA 91125, USA. Email: `nayak@cs.caltech.edu`.

[†] AT&T Labs–Research, 180 Park Ave, Florham Park, NJ 07932, USA. Email: `shor@research.att.com`.

Abstract

In this paper, we focus on a special framework for quantum coin flipping protocols, *bit-commitment based protocols*, within which almost all known protocols fit. We show a lower bound of 1/16 for the bias in any such protocol. We also analyse a sequence of multi-round protocol that tries to overcome the drawbacks of the previously proposed protocols, in order to lower the bias. We show an intricate cheating strategy for this sequence, which leads to a bias of 1/4. This indicates that a bias of 1/4 might be optimal in such protocols, and also demonstrates that a cleverer proof technique may be required to show this optimality.

1 Quantum coin flipping

Coin flipping is the communication problem in which two distrustful parties wish to agree on a common random bit, by "talking over the phone" [1]. When the two parties follow a protocol honestly, the bit they agree on is required to be 0 or 1 with equal probability. Ideally, they would also like that if any (dishonest) party deviates from the protocol, they do not agree on any particular outcome with probability more than 1/2. It is known that ideal coin flipping is impossible, in both, the classical and the quantum setting [2, 3]. In fact, in any classical protocol, one of the two parties can force the outcome of the protocol to a value of her choice with probability 1. In [4], Aharonov, Ta-Shma, Vazirani, and Yao showed that it is possible to design a *quantum* coin flipping protocol in which no player can force the outcome of the protocol with probability more than a constant $1/2 + \epsilon$, with bias a constant $\epsilon < 1/2$. In other words, any cheating player in such protocols is detected with constant probability. Later, Ambainis [5] gave an improved protocol with bias at most 1/4.

Formally, a quantum coin flipping protocol with bias ϵ is a two-party communication game in the style of [6], in which the players start with no inputs, and compute values $c_A, c_B \in \{0, 1\}$ respectively (or declare that the other player is cheating). The protocol satisfies the following additional properties:

1. If both players are honest (i.e., follow the protocol), then they agree on the outcome of the protocol: $c_A = c_B$, and the outcome is 0 or 1 with equal probability: $\Pr(c_A = c_B = b) = 1/2$, for $b \in \{0, 1\}$.

2. If one of the players is honest (i.e., the other player may deviate arbitrarily from the protocol in his or her local computation), then the outcome of the protocol has bias at most ϵ: for any $b \in \{0, 1\}$, $\Pr(c_A = c_B = b) \leq 1/2 + \epsilon$.

Almost all quantum coin flipping protocols with bias that is provably smaller than a half [4, 5] are based on the notion of bit-commitment. In other words, they have the following form, when the parties flipping the coin, Alice and Bob, are honest.[1]

Protocol schema Π:

1. First, Alice and Bob each pick a random bit, a and b respectively, and privately construct states ρ_a and σ_b. The states are over three sets of qubits, a *commitment* part, a *revelation* part, and a *verification* part. The revelation part consists of one qubit that contains the value of the bit picked.

2. Next, they *commit* to their respective bits a and b, by sending each other the commitment part of their states ρ_a and σ_b. They may do this over several rounds of communication, in which they send messages alternately.

3. Then, they *reveal* to the other party the bits a, b they picked (in some order), and follow that up by sending the rest of the states ρ_a, σ_b (the verification part). This may again be over several rounds of communication. This allows the each party to check via suitable measurements that the state with which the other, Alice (or Bob), committed to her (his) bit is indeed consistent with a (b).

The result of the protocol is $c = a \oplus b$, if neither player is detected cheating during the third (verification) stage.

[1] We were recently informed [7] of a protocol of a different kind that also achieves a bias of 1/4.

For example, in the case of the protocol in [5], Alice uses the right half of the following state to commit to her bit a, and the left half to help Bob check her commitment:

$$\begin{aligned} \rho_a &= \sum_{s=0,1} \frac{1}{2} |a,s\rangle\langle a,s| \otimes |\psi_{a,s}\rangle\langle\psi_{a,s}|, \qquad \text{where} \\ |\psi_{x,s}\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + (-1)^s |x+1\rangle). \end{aligned}$$

Bob skips the commit stage, and directly reveals his bit b. In the final stage, Bob checks that the state of the qutrit Alice sent in the first round is indeed consistent with a, s, by measuring it in a basis containing $|\psi_{a,s}\rangle$.

Protocols of the form described above may be recast in the following terms. First, Alice and Bob each pick a random bit. Then, they successively send each other qubits *which do not depend on the qubits sent by the other party in the previous rounds*. The qubits sent by Alice represent a commitment to her bit a along with auxiliary information required by Bob to check if she is cheating. Similarly for the qubits sent by Bob. Thus, after all communication is over, the states Alice and Bob sent to each other for 0 and 1 are perfectly distinguishable. They measure the states received from the other party (possibly with some ancilla) according to a von Neumann measurement to determine the bits a, b or to detect cheating. The outcome of the coin flip is $a \oplus b$ if no cheating is detected. In this description, we have assumed, w.l.o.g. [8], that all measurements are done at the end. Note that the description is also slightly more general in that the players may not explicitly reveal the bits they intend to commit to, and the commitment and the verification stages may be interleaved. We will henceforth refer to such protocols as *bit-commitment based protocols*.

In this paper, we study coin flipping protocols that fall into the special framework described above, that of bit-commitment based protocols. We show a lower bound of $1/16$ for the bias in any such protocol (Theorem 2.2). This provides a single proof that these protocols, including the one proposed in [9], cannot lead to arbitrarily small bias.

Next, we analyse a sequence of protocols that tries to overcome the drawbacks of the previously proposed protocols, and also tries to circumvent the cheating strategy that leads to this above lower bound. We show an intricate cheating strategy for the sequence of protocols, which leads to a bias of $1/4$. This indicates that a bias of $1/4$ might be optimal in such protocols, and also demonstrates that a cleverer proof technique than the one used in Theorem 2.2 is required to show this optimality.

Recently, Kitaev has shown a lower bound of $1/\sqrt{2} - 1/2 = 0.2071\ldots$ for the bias in *arbitrary* quantum coin flipping protocols [10]. This is bigger than the bias of $1/16 = 0.0625$ that we show (and it applies to *any* coin-flipping protocol). In more detail, Kitaev's result states that the product of the optimum cheating probabilities for the two parties is at least $1/2$. We leave it as an open question whether these optima *sum* to at least $3/2$, as is the case in all known protocols, including the ones considered in this paper.

Kitaev's $1/\sqrt{2} - 1/2$ lower bound doesn't seem to apply to quantum games in which the two parties involved compete to "win" by getting a particular value of the coin as the outcome (say Alice wins if the outcome is 0, and Bob wins if the outcome is 1). This is also known as *weak* coin-flipping in the literature. Protocols for weak coin-flipping with bias less than $1/4$ have been discovered. Kerenidis and Nayak [11] have shown a protocol with cheating probability at most $0.739\ldots$. Ambainis [12], and Spekkens and Rudolph [13] have shown protocols with an even lower cheating probability of at most $1/\sqrt{2} = 0.707\ldots$. No lower bound is known for the bias in weak coin flipping; arbitrarily small bias may still be possible.

2 A lower bound on the bias

We first show that any bit-commitment based protocol may be reduced to an extremely simple protocol of the same type, with bias at most that in the original protocol, and by increasing the number of rounds by at most 1.

Lemma 2.1 *For any bit-commitment based coin flipping protocol $\mathcal{P}$ (of the form Π, or more generally, as described in Section 1), there is another such protocol $\mathcal{P}'$ such that*

1. *The states ρ_a, σ_b are pure: $\rho_a = |\psi_a\rangle\langle\psi_a|$ and $\sigma_b = |\phi_b\rangle\langle\phi_b|$, where $|\psi_0\rangle \perp |\psi_1\rangle$ and $|\phi_0\rangle \perp |\phi_1\rangle$,*

2. *Alice measures the state she received from Bob according to the measurement given by the operators $P_0 = |\phi_0\rangle\langle\phi_0|$, $P_1 = |\phi_1\rangle\langle\phi_1|$, and $P_{\mathrm{c}} = I - P_0 - P_1$, to determine Bob's bit or to check if he is cheating. Bob does an analogous measurement given by Q_0, Q_1, Q_{c} on the state he receives from Alice, and*

3. *The bias is at most the bias of $\mathcal{P}$.*

Proof: The protocol $\mathcal{P}'$ is obtained by stipulating that the players use a fixed *purification* $|\psi_a\rangle, |\phi_b\rangle$ of the states ρ_a, σ_b used in the original protocol $\mathcal{P}$. Since the states ρ_0, ρ_1 are perfectly distinguishable, their purifications are orthogonal. Similarly with ϕ_0, ϕ_1. All but the last two rounds of $\mathcal{P}'$ are as in the original one. We stipulate that the players send the entire (purified) state ψ_a or ϕ_b in $\mathcal{P}'$. Thus, the last player to send a message in $\mathcal{P}$ sends the qubits used in the purification in the penultimate round of $\mathcal{P}'$. In the final round, the other player sends the qubits used in purifying her state. We also alter the measurement to the ones mentioned in the lemma.

We now show that this modification of the protocol results in bias at most that in the original one. We do this by showing that any cheating strategy of a player in the modified protocol $\mathcal{P}'$ that achieves a bias of ϵ leads to a cheating strategy in the original protocol $\mathcal{P}$ with at least the same bias.

For concreteness, we consider a cheating strategy for Alice in the protocol $\mathcal{P}'$. (The argument for the case of Bob is similar.) In her strategy in $\mathcal{P}$, Alice acts exactly as in the original strategy, except that she is not required to send the "purification qubits" meant for her last message in $\mathcal{P}'$. We need only show that the probability with which the measurement in $\mathcal{P}$ yields 0 or 1 dominates the same probability for $\mathcal{P}'$.

Suppose that Bob uses a von Neumann measurement given by the projection operators R_0, R_1, R_{c} in $\mathcal{P}$. We concentrate on the probability that Alice can convince Bob that she had picked $a = 0$. The other probability may be bounded similarly. Since $R_0 \rho_0 R_0 = \rho_0$, the purification $|\psi_0\rangle$ lies in the range of $I \otimes R_0$. The states of the qubits sent to Bob by Alice and her private qubits in $\mathcal{P}'$ are together given by some mixed state $\sum_j \mu_j |\xi_j\rangle\langle\xi_j|$ (where the states $|\xi_j\rangle$ are over the space of ρ_a and the purification space). It thus suffices to show that for any state $|\xi\rangle \in \{|\xi_j\rangle\}$,

$$\| \, I \otimes R_0 |\xi\rangle \, \|^2 \quad \geq \quad |\langle\psi_0|\xi\rangle|^2 .$$

Note that the LHS is the squared-norm of the projection of $|\xi\rangle$ onto the range of $I \otimes R_0$, and the RHS is the squared-norm of the projection of the same vector onto a *subspace* of that range, the one-dimensional space spanned by $|\psi_0\rangle$. The inequality is then immediate. This shows that Alice can achieve at least the same bias as in $\mathcal{P}'$. ■

This simple characterisation of bit-commitment based protocols proves useful in the analysis of the smallest bias achievable with such protocols. Using this, we show that coin flipping protocols based on bit-commitment cannot achieve arbitrarily small bias.

Theorem 2.2 *In any quantum coin flipping protocol based on bit-commitment, one of the parties can achieve probability of cheating at least* $9/16$*.*

Proof: As shown in Lemma 2.1, any such protocol between honest parties may be viewed as follows: first, Alice and Bob construct the states $|\psi_a\rangle$ and $|\phi_b\rangle$, respectively, corresponding to the random bits a and b. Then, they send each other a part of the states $|\psi_a\rangle, |\phi_b\rangle$ a few qubits at a time. Finally, they measure the qubits received from each other using projections $P_0, P_1, P_{\rm c}$ and $Q_0, Q_1, Q_{\rm c}$.

Let $\rho_{a,i} = \mathrm{Tr}_{A_i}(|\psi_a\rangle\langle\psi_a|)$ be the state sent to Bob by Alice by round i (so A_i are the qubits of $|\psi_a\rangle$ still with Alice after the i-th round). Let $\sigma_{b,i}$ be the corresponding state sent to Alice by Bob.

Let there be n rounds in all. Let $F_{A,i} = F(\rho_{0,i}, \rho_{1,i})$ and similarly $F_{B,i} = F(\sigma_{0,i}, \sigma_{1,i})$. Here, $F(\cdot,\cdot)$ is *fidelity* function as defined in [14]. So $F_{A,0} = F_{B,0} = 1$. Note that $Q_a \rho_{a,n} Q_a = \rho_{a,n}$, and similarly $P_b \sigma_{b,n} P_b = \sigma_{b,n}$, so that $F_{A,n} = F_{B,n} = 0$.

Lemma 2.3 *Consider a protocol with honest players. For any constant* $0 \leq \alpha \leq 1$*, there is a player, say Alice, and a round* $k \geq 0$ *such that the states she sends to Bob by the* k*-th round on* $a = 0$ *and* 1 *have fidelity at least* α*, and the fidelity of the states she receives from Bob by the next round have fidelity at most* α*. In other words,* $F_{A,k} \geq \alpha$ *and* $F_{B,k+1} \leq \alpha$*.*

Proof: The case $\alpha = 1$ is trivial. Let $\alpha < 1$. Note that both $F_{A,i}, F_{B,i}$ decrease from 1 to 0 through the course of the protocol. Consider the first round $i \geq 1$ such that one of these, say $F_{B,i}$, becomes $\leq \alpha$. Round $k = i - 1$ satisfies the property we seek. ■

We will devise a cheating strategy for a player as given by Lemma 2.3 above for $\alpha = 1/4$. Say this player is Alice, and the round identified in the lemma is k. There is a unitary transformation on the qubits A_k which achieves maximum fidelity between $\rho_{0,k}, \rho_{1,k}$ [14], i.e.,

$$|\langle\psi_0|U|\psi_1\rangle|^2 = F_{A,k} = F(\rho_{0,k}, \rho_{1,k}) \geq \frac{1}{4}.$$

Moreover, we may assume that $\langle\psi_0|U|\psi_1\rangle$ is real and non-negative.

Alice may cheat as follows. She constructs the state

$$|\xi\rangle = \frac{|\psi_0\rangle + U|\psi_1\rangle}{\parallel \psi_0 + U\psi_1 \parallel}$$

and uses this state in the protocol till round $k+1$. After this round, she makes the best possible measurement to distinguish $\sigma_{0,k+1}$ and $\sigma_{1,k+1}$ to guess the value of b. If her guess g is 0, she proceeds with the rest of the protocol, Otherwise, she applies $U^\dagger$ to her part of $|\xi\rangle$ (the qubits in A_k) and then completes the protocol.

Lemma 2.4 $\Pr(c = 0 | g = b) \geq \frac{1+\sqrt{F_{A,k}}}{2}$.

Proof: First, note that

$$\| \psi_0 + U\psi_1 \|^2 \quad = \quad \| \psi_0 \|^2 + \| \psi_1 \|^2 + 2\langle \psi_0 | U | \psi_1 \rangle \quad = \quad 2(1 + \sqrt{F_{A,k}}).$$

Suppose $b = 0$. (The other case is similar.) Note that the probability that Alice succeeds in getting the outcome $c = 0$ given that she guesses the value of b correctly is

$$\| Q_0 \xi \|^2 \quad = \quad \frac{\| Q_0\psi_0 + Q_0 U \psi_1 \|^2}{\| \psi_0 + U\psi_1 \|^2} \quad = \quad \frac{\| Q_0\psi_0 \|^2 + \| Q_0 U\psi_1 \|^2 + 2\langle \psi_0 | Q_0 U | \psi_1 \rangle}{2(1 + \sqrt{F_{A,k}})}.$$

Now $Q_0|\psi_0\rangle = |\psi_0\rangle$, since $|\psi_0\rangle$ is the state Alice would have used if she were honest. So the first term in the numerator above is 1, and the last term is $2\sqrt{F_{A,k}}$. The second term may be bounded from below by noting that since $|\psi_0\rangle$ belongs to the range of Q_0, the square-norm of the projection $Q_0 U|\psi_1\rangle$ is at least $\langle \psi_0|U|\psi_1\rangle^2 = F_{A,k}$. Thus, the probability of cheating is at least

$$\frac{1 + F_{A,k} + 2\sqrt{F_{A,k}}}{2(1 + \sqrt{F_{A,k}})} \quad = \quad \frac{1 + \sqrt{F_{A,k}}}{2},$$

which is the bound claimed. ■

The probability that Alice correctly guesses b is (using Bayes' strategy)

$$\Pr(g = b) \quad = \quad \frac{1}{2} + \frac{\| \sigma_{0,k+1} - \sigma_{1,k+1} \|_{\mathrm{tr}}}{4}.$$

By a result of Fuchs and van de Graaf [15],

$$\| \sigma_{0,k+1} - \sigma_{1,k+1} \|_{\mathrm{tr}} \quad \geq \quad 2(1 - \sqrt{F_{B,k+1}}) \quad \geq \quad 2(1 - \frac{1}{2}) \quad = \quad 1.$$

The net probability that Alice succeeds in biasing the coin towards 0 is therefore

$$\Pr(c = 0) \quad \geq \quad \Pr(c = 0 | g = b) \cdot \Pr(g = b) \quad \geq \quad \frac{3}{4} \cdot \frac{3}{4} \quad = \quad \frac{9}{16}.$$

This proves the theorem. ■

3 A sequence of highly interactive protocols

In this section we look at a sequence of bit-commitment based protocols in which Alice and Bob very gradually send each other information about their bits in the commit stage. Intuitively, such protocols seem to be good candidates for achieving bias much smaller than 1/4, since a dishonest player does not get much information about the other's bit until a significant number of rounds have elapsed, and he would have heavily committed to some bit by then. However, this intuition appears to be mistaken, and we give intricate cheating strategies for each of the players with which at least one of them can achieve bias at least as high as 1/4. This suggests that the optimal bias for this kind of protocol might be 1/4, and also that proving this optimality might require ideas more sophisticated than those in Theorem 2.2.

Define, for $x, s \in \{0, 1\}$,

$$|\psi(x, s)\rangle \quad = \quad \sqrt{1 - \epsilon}\,|0\rangle + (-1)^s \sqrt{\epsilon}\,|x + 1\rangle.$$

These states provide the best trade-off between how much information they reveal, and how much cheating in commitment they allow.

The protocol $\mathcal{P}_n$, $n \in \{1, 2, 3, \ldots\}$, goes as follows. Alice picks $a \in_{\mathrm{R}} \{0, 1\}$, and Bob picks $b \in_{\mathrm{R}} \{0, 1\}$. Then they alternately send each other, for a total of n rounds, the states $|\psi(a, s)\rangle$ and $|\psi(b, s)\rangle$ respectively, for independently chosen random sign s, starting with Alice. The last player to receive such a state then reveals the bit he/she chose, followed by the other player. Then, they reveal the signs used in their states in the opposite order, and the other party checks the state he/she received against the claimed bit and sign. If no cheating is detected, the players declare the $c = a \oplus b$ as the result of the protocol.

More formally,

1. For $i = 1, 2, 3, \ldots, n$, if i is odd, Alice picks $s_i \in_{\mathrm{R}} \{0, 1\}$, and sends $|\psi(a, s_i)\rangle$ to Bob, else, if i is even, Bob picks $s_i \in_{\mathrm{R}} \{0, 1\}$, and sends $|\psi(b, s_i)\rangle$ to Alice.

2. If n is even, Alice sends a to Bob, and then Bob sends b to Alice. Otherwise, if n is even they reveal their bits in the opposite order.

3. For $i = n, n-1, n-2, \ldots, 1$, the player that picked s_i reveals it to the other player. In other words, the "signs" used in the states are revealed in the opposite order: If i is odd, Alice sends s_i to Bob, else Bob sends s_i to Alice. The player that receives this bit checks via a measurement that the state sent to her/him in the i-th round of the protocol is indeed consistent with the bit and the sign that the other player sent.

   Note that the order of revealing signs is designed so that the naïve strategy of reusing a state that a player got in a previous round does not work.

4. If *all* the checks are passed, the outcome of the protocol is $c = a \oplus b$.

Before we give cheating strategies for these protocols, we analyse general versions of $\mathcal{P}_1$ and $\mathcal{P}_2$. This illustrates the main approach taken in the strategies for $\mathcal{P}_n$, $n \geq 3$.

3.1 The three-round version

We start by analysing the protocol $\mathcal{P}_1$ with one round of commitment. This happens to be a parametrised version of the three-round protocol due to Ambainis [5]. We prove a property of this protocol that helps us analyse protocols with more rounds.

The protocol may be described with a parameter $\alpha \in [0, \pi]$ (such that $\tan\frac{\alpha}{2} = \sqrt{\frac{\epsilon}{1-\epsilon}}$) as:

1. Alice picks $a, s \in_{\mathrm{R}} \{0, 1\}$, and sends Bob the state $|\psi_{a,s}\rangle$, where the state is defined as follows:

$$|\psi_{a,s}\rangle \quad = \quad \cos\frac{\alpha}{2}|0\rangle + (-1)^s \sin\frac{\alpha}{2}|a+1\rangle. \tag{1}$$

2. Bob picks $b \in_{\mathrm{R}} \{0, 1\}$ and sends it to Alice.

3. Alice then reveals the bits a, s to Bob, who checks for consistency with the state initially sent by Alice.

The output of the protocol is given by $c = a \oplus b$, if all the checks are passed.

Lemma 3.1 *If Bob is honest, then* $\Pr(c=0) \leq \frac{1}{4}(3+\cos\alpha)$.

Proof: The analysis proceeds as in [5], by symmetrising the strategy of a dishonest Alice, so that in the last round, she sends $s=0$ and $s=1$ with equal probability, and assuming that she always sends $a=b$ (this only increases her chances of cheating successfully). For such a symmetric strategy, Ambainis [5, Lemma 8] shows that

$$\Pr(c=0) \quad \leq \quad \frac{F(\rho_0,\rho)+F(\rho_1,\rho)}{2}, \tag{2}$$

where ρ is the state of Bob after the first round, and ρ_0,ρ_1 are the analogous states corresponding to $b=0,1$ respectively, if Alice were honest:

$$\rho_a \quad = \quad \frac{1}{2}(|\psi_{a,0}\rangle\langle\psi_{a,0}| + |\psi_{a,1}\rangle\langle\psi_{a,1}|),$$

and $F(\sigma_0,\sigma_1) = \|\sqrt{\sigma_0}\sqrt{\sigma_1}\|_{\mathrm{tr}}^2$ denotes the fidelity of two density matrices [14].

We show in Lemma 3.2 below that the expression in equation (2) is bounded above by

$$\frac{1}{2}(1+F(\rho_0,\rho_1)^{1/2}).$$

Since $F(\rho_0,\rho_1) = \cos^4\frac{\alpha}{2}$, the bound follows. ■

We now prove the lemma mentioned above.

Lemma 3.2 *For any two density matrices* σ_0,σ_1,

$$\max_{\sigma} \quad F(\sigma_0,\sigma)+F(\sigma_1,\sigma) \quad \leq \quad 1+F(\sigma_0,\sigma_1)^{1/2}.$$

Proof: Let σ be the density matrix that achieves the maximum, and let $|\phi_x\rangle$ be a purification of σ_x, for $x=0,1$. Let $|\xi_x\rangle$ be the purification of σ that achieves maximum fidelity with σ_x [14]:

$$F(\sigma_0,\sigma) \quad = \quad |\langle\phi_x|\xi_x\rangle|^2.$$

Since $|\xi_0\rangle, |\xi_1\rangle$ are the purifications of the same density matrix σ, there is a local unitary operator U such that $|\xi_1\rangle = U|\xi_0\rangle$. Now,

$$\begin{aligned}
F(\sigma_0,\sigma)+F(\sigma_1,\sigma) &= |\langle\phi_0|\xi_0\rangle|^2 + |\langle\phi_1|\xi_1\rangle|^2 \\
&= |\langle\phi_0|\xi_0\rangle|^2 + |\langle\phi_1|U|\xi_0\rangle|^2 \\
&\leq \max_{|\xi\rangle} |\langle\phi_0|\xi\rangle|^2 + |\langle\phi_1|U|\xi\rangle|^2 \\
&= 1+|\langle\phi_1|U|\phi_0\rangle| \qquad (3)\\
&\leq 1+\max_{\text{local } U} |\langle\phi_1|U|\phi_0\rangle| \\
&= 1+F(\sigma_0,\sigma_1)^{1/2}. \qquad (4)
\end{aligned}$$

Equation (3) above follows by noticing that the state $|\xi\rangle$ that achieves the maximum is the vector that bisects the angle between $|\phi_0\rangle$ and $U|\phi_1\rangle$. Another way of getting the bound is by noticing that the expression is the maximum eigenvalue of the matrix $|\phi_0\rangle\langle\phi_0| + U|\phi_1\rangle\langle\phi_1|U^\dagger$. The last step—equation (4)—follows from a characterisation of fidelity due to Jozsa [14]. ■

3.2 The five-round protocol

Next, we give a tight analysis for a general five-round version of the protocol $\mathcal{P}_2$ described above. This version of the protocol still does not improve over the bias of 1/4 achieved by the three-round protocol of [5]. However, it suggests a better cheating strategy for the many-rounds version than the one given in the proof of Theorem 2.2.

The version of the protocol $\mathcal{P}_2$ we consider has the following five rounds with honest players:

1. Alice picks $a, s \in_{\mathrm{R}} \{0, 1\}$, and sends Bob the state $|\psi_{a,s}\rangle$, where the state is defined as follows:

$$|\psi_{a,s}\rangle \quad = \quad \cos\frac{\alpha}{2}|0\rangle + (-1)^s \sin\frac{\alpha}{2}|a+1\rangle, \tag{5}$$

   for an angle α to be specified later.

2. Similarly, Bob picks $b, s' \in_{\mathrm{R}} \{0, 1\}$ and sends Alice the state $|\phi_{b,s'}\rangle$, where the state is defined as follows:

$$|\phi_{b,s'}\rangle \quad = \quad \cos\frac{\beta}{2}|0\rangle + (-1)^{s'} \sin\frac{\beta}{2}|b+1\rangle, \tag{6}$$

   for an angle β to be specified later.

3. Alice then reveals the bit a to Bob.

4. Bob reveals both b and s' to Alice, and Alice verifies, by an appropriate measurement, that the state sent by Bob is consistent with b, s'.

5. Alice now discloses s as well, and Bob verifies that the state sent by Alice in the first round is consistent with a, s.

The output of the protocol, the coin flip c, is given by the exclusive-or of two bits, $c = a \oplus b$ (provided no cheating is detected).

In the case the players are honest, $\Pr(c = 0) = \Pr(c = 1) = 1/2$. Below we prove an upper bound on the probability that any player can achieve (by deviating from the protocol) for an outcome of their choice. In the following discussion, we will assume, w.l.o.g., that a dishonest player prefers the outcome $c = 0$.

First, we prove a bound on the probability that a dishonest Bob can achieve, provided Alice is honest.

Lemma 3.3 *If Alice is honest, then* $\Pr(c = 0) \ \leq \ 1 - \frac{1}{2}\cos^2\frac{\alpha}{2}\sin^2\frac{\beta}{2}$.

Proof: We claim that Bob's optimal cheating strategy is to measure the state received from Alice in the standard basis, and then commit to a state according to the outcome. If he observes $|1\rangle$ or $|2\rangle$, he can cheat with probability 1 in the rest of the protocol. In case he observes $|0\rangle$, his probability of cheating successfully is bounded by a constant less than one. This gives us the bound, as explained below.

Note that the last round is of no consequence to Bob's cheating strategy, and we may trace the sign bit s out after Alice sends the first message (and eliminate the last round). The protocol then becomes equivalent to one in which the first round takes the following form:

> Alice picks $b \in_{\mathrm{R}} \{0, 1\}$, and sends $|0\rangle$ with probability $\cos^2\frac{\alpha}{2}$, and $|b+1\rangle$ with probability $\sin^2\frac{\alpha}{2}$.

This reduces the protocol to a convex combination of two protocols, with weights $\cos^2\frac{\alpha}{2}$ and $\sin^2\frac{\alpha}{2}$. In the first protocol, the first message is the same regardless of the value of b, and in the second, the first message reveals b entirely.

In the first case, the protocol reduces to a three-round protocol of the type studied in Section 3.1, with the role of Alice and Bob reversed. We may now use Lemma 3.1 to bound the probability that Bob can cheat by $(3+\cos\beta)/4$.

In the second case, the protocol becomes trivial, and Bob can cheat with probability 1.

The probability of convincing Alice that $c = 0$ is thus bounded by

$$\frac{3+\cos\beta}{4}\cos^2\frac{\alpha}{2} \;+\; \sin^2\frac{\alpha}{2},$$

which reduces to the bound we seek. ■

Bob can easily achieve the probability bound stated in the lemma, by following a cheating strategy as in [5, Lemma 10]. This shows that the analysis in our proof is optimal.

We now turn to the case where Alice is dishonest. The following lemma bounds Alice's cheating probability.

Lemma 3.4 *If Bob is honest, then* $\Pr(c=0) \quad \leq \quad \frac{1}{2}(1+\cos^2\frac{\alpha}{2}\sin^2\frac{\beta}{2})$.

Proof: If Alice is dishonest, she tries to force the outcome $c = 0$ by guessing the bit that Bob picked, and then convincing him that the state she sent is consistent with that guess.

We begin by symmetrising Alice's strategy so that the density matrix of the qutrit she commits to in the first round is diagonal in the $0, 1, 2$ basis. This is done in a manner as in in [5, Lemma 6]; we omit the details. We may therefore assume, w.l.o.g., that the joint state of Alice and Bob after the communication in the first round is

$$|\xi\rangle \quad = \quad \sqrt{\lambda_0}\,|0,0\rangle + \sqrt{\lambda_1}\,|1,1\rangle + \sqrt{\lambda_2}\,|2,2\rangle,$$

where $\sum_i \lambda_i = 1$ and the density matrix of Bob's part is

$$\rho \quad = \quad \lambda_0\,|0\rangle\langle 0| + \lambda_1\,|1\rangle\langle 1| + \lambda_2\,|2\rangle\langle 2|.$$

Alice's strategy in the third round is given by some measurement on her entangled qutrit, the qutrit sent by Bob, and some ancilla. This measurement gives the value of the bit she sends to Bob in that round. We consider the superoperator T acting on their joint state consisting of this measurement, composed with the tracing out of all subsystems except the bit sent in that round, and the qutrit sent in the first round. This superoperator is determined entirely by its action on $|i\rangle\langle i| \otimes |\xi\rangle\langle\xi|$, for $i = 0, 1, 2$, since the qutrit sent by Bob is a mixture of the states $|i\rangle\langle i|$. Let

$$T(|i\rangle\langle i| \otimes |\xi\rangle\langle\xi|) \quad = \quad |0\rangle\langle 0| \otimes \rho_{i0} + |1\rangle\langle 1| \otimes \rho_{i1},$$

where $\sum_j \rho_{ij} = \rho$. Thus, the unnormalised density matrix of the qutrit Alice sent in the first round is:

$$\cos^2\frac{\beta}{2}\,\rho_{0b} + \sin^2\frac{\beta}{2}\,\rho_{b+1,b}, \tag{7}$$

given that Bob had picked bit b, and Alice's guessed it correctly: $a = b$.

Let $\tilde{\rho}_{ij}$ be the state, diagonal in the $0,1,2$ basis, obtained by measuring the state ρ_{ij}, and let:

$$\begin{aligned}\tilde{\rho}_{00} &= \mathrm{diag}(\mu_0, \mu_1, \lambda_2 - \mu_2)\\ \tilde{\rho}_{01} &= \mathrm{diag}(\lambda_0 - \mu_0, \lambda_1 - \mu_1, \mu_2),\end{aligned}$$

where $0 \leq \mu_i \leq \lambda_i$. Finally, let $\tilde{\rho}_b = \cos^2 \frac{\beta}{2}\, \tilde{\rho}_{0b} + \sin^2 \frac{\beta}{2}\, \tilde{\rho}_{b+1,b}$ be the state (7) measured in the $0,1,2$ basis. Let σ_b be the density matrix of the qutrit Alice would send in the first round, if she were honest, and had picked the bit b:

$$\sigma_b = \cos^2 \frac{\alpha}{2}\, |0\rangle\langle 0| + \sin^2 \frac{\alpha}{2}\, |b+1\rangle\langle b+1|.$$

Assume now that instead of having picked s' at random, Bob had created a uniform superposition over the two possible values for the sign bit, and that he sends the qubit containing s' in the fourth round. It is now not hard to prove (cf. [5]) that the probability that (the dishonest) Alice is able to convince Bob that the state she sent is consistent with b, is at most $F(\tilde{\rho}_b, \sigma_b)$. This probability is thus also bounded by $F(\tilde{\sigma}_b, \sigma_b)$, where

$$\begin{aligned}\tilde{\sigma}_0 &= \cos^2 \frac{\beta}{2}\, (\mu_0\, |0\rangle\langle 0| + \lambda_1\, |1\rangle\langle 1|) \;+\; \sin^2 \frac{\beta}{2}\, \tilde{\rho}\\ \tilde{\sigma}_1 &= \cos^2 \frac{\beta}{2}\, ((\lambda_0 - \mu_0)\, |0\rangle\langle 0| + \lambda_2\, |2\rangle\langle 2|) \;+\; \sin^2 \frac{\beta}{2}\, \tilde{\rho}.\end{aligned}$$

In other words, it only helps Alice to claim that she sent a state corresponding to bit b if she either sees $b+1$ in the qutrit she receives from Bob, or if she sees 0 and had sent the state $|b+1\rangle$ to Bob in the first round.

By optimising over the choice of μ_0 for fixed λ_i and then optimising λ_i, we see that the optimum of $F(\tilde{\sigma}_0, \sigma_0) + F(\tilde{\sigma}_1, \sigma_1)$ is achieved when

$$\begin{aligned}\mu_0 &= \frac{\lambda_0(\lambda_1 - \lambda_2 \sin^2 \frac{\beta}{2})}{(\lambda_1 + \lambda_2) \cos^2 \frac{\beta}{2}},\\ \lambda_1 &\geq \lambda_2\, \sin^2 \frac{\beta}{2}, \quad \text{and}\\ \lambda_2 &\geq \lambda_1\, \sin^2 \frac{\beta}{2}.\end{aligned}$$

We thus get the following bound on Alice's cheating probability:

$$\frac{1}{2}(1 + \cos^2 \frac{\alpha}{2} \sin^2 \frac{\beta}{2}),$$

which is the bound claimed. ∎

Next, we describe a cheating strategy for Alice that achieves the outcome of her choice with probability as high as in the upper bound above, showing that our analysis is optimal.

Lemma 3.5 *If Bob is honest, Alice can achieve* $\Pr(c = 0) \geq \frac{1}{2}(1 + \cos^2 \frac{\alpha}{2} \sin^2 \frac{\beta}{2})$.

Proof: Alice constructs the following entangled state and sends one half of it to Bob in the first round:

$$\begin{aligned}|\xi\rangle &= \sqrt{1-\lambda}|0,0\rangle + \sqrt{\frac{\lambda}{2}}|1,1\rangle + \sqrt{\frac{\lambda}{2}}|2,2\rangle, \quad \text{where}\\ \lambda &= \frac{\sin^2 \frac{\alpha}{2}}{(1 + \sin^2 \frac{\beta}{2}) \cos^2 \frac{\alpha}{2} + \sin^2 \frac{\alpha}{2}},\end{aligned}$$

and so the density matrix of the qutrit with Bob is $\rho = \mathrm{diag}(1-\lambda, \lambda/2, \lambda/2)$ in the $0, 1, 2$ basis.

After Alice receives a qutrit from Bob in the second round, she applies a unitary transformation to "guess" a value for b to maximise her chances of getting $c = 0$. This transformation acts on the qutrit she received from Bob, her entangled qutrit from the first round, and an ancilla qubit. It can be written as:

$$|0\rangle\langle 0| \otimes (|0\rangle\langle 0| \otimes H + |1\rangle\langle 1| \otimes I + |2\rangle\langle 2| \otimes \sigma_{\mathrm{x}}) \quad + \quad |1\rangle\langle 1| \otimes I \otimes I \quad + \quad |2\rangle\langle 2| \otimes I \otimes \sigma_{\mathrm{x}}.$$

where H is the Hadamard transform, and σ_{x} is the Pauli "bit flip" matrix. In other words, Alice guesses $a = b$, if the qutrit from Bob reveals the identity of b. Otherwise, if she committed with a $|d\rangle$, $(d = 1, 2)$ she says $a = d - 1$, since this commitment is irrevocable. If she committed with $|0\rangle$, she says $a = 0, 1$ with equal amplitude. It is crucial that she does this *in superposition*.

In the fourth round, Bob reveals the state of the qutrit he sent. If Alice sees that the bit Bob picked is different from her guess, i.e. $b \neq a$, then she sends, say, $s = 0$ (an arbitrary value, since she has lost the game). Otherwise, she tries to pick the best s possible to maximise her chance of passing Bob's check. We describe her actions when $b = 0, s' = 0$; the other cases are similar.

The unnormalised density matrix of the qutrit she sent in the first round, when $b = 0, s' = 0$, conditioned on her reply being $a = 0$ is:

$$\cos^2\frac{\beta}{2}\left(\frac{1-\lambda}{2}\,|0\rangle\langle 0| + \frac{\lambda}{2}\,|1\rangle\langle 1|\right) \quad + \quad \sin^2\frac{\beta}{2}\,\rho.$$

Note that she knows the state of all their qubits, given b, s', a. She can thus transform her part of the state so that their joint state looks like

$$\sqrt{\frac{1-\lambda}{2}(1+\sin^2\frac{\beta}{2})}\,|00\rangle \quad + \quad \sqrt{\frac{\lambda}{2}}\,|11\rangle \quad + \quad \sqrt{\frac{\lambda}{2}}\,\sin\frac{\beta}{2}\,|22\rangle.$$

She sends $s = 0$ if the entangled bit is 2. If her entangled qutrit is not 2, she does a Hadamard transform on her entangled qubit, and sends that to Bob as s. The probability with which Bob accepts $a = b, s$ is then

$$\left(\sqrt{\frac{1-\lambda}{2}(1+\sin^2\frac{\beta}{2})}\,\cos\frac{\alpha}{2} + \sqrt{\frac{\lambda}{2}}\,\sin\frac{\alpha}{2}\right)^2,$$

which evaluates to the expression stated in the lemma. She can achieve the same probability of success for all other values of b, s' as well. Thus, the overall chance of her succeeding in cheating is also given by this expression. ∎

The properties we established above show that this protocol still has a bias of $1/4$: the cheating probability for Alice and Bob are of the form $\frac{1}{2}+\delta$ and $1-\delta$ respectively (with $\delta = \frac{1}{2}\cos^2\frac{\alpha}{2}\sin^2\frac{\beta}{2}$), and their maximum is minimised when $\delta = 1/4$.

3.3 Cheating strategies for $\mathcal{P}_n$

Let $A_n(\epsilon), B_n(\epsilon)$ be the maximum value of Alice and Bob's cheating probability in the protocol $\mathcal{P}_n$, when the other party is honest.

Bob's optimal strategy may be reduced to a strategy for Alice as in the proof of Lemma 3.3. Alice always starts the protocol, and also sends the last message. Since the last message does not affect Bob's cheating strategy, we may trace it out of the protocol when analysing his optimal strategy. The protocol $\mathcal{P}_n$ then reduces to a mixture of protocols where Alice sends $|0\rangle$ in the first round with probability $1-\epsilon$ (i.e. does not reveal any information about the bit a), and sends $|a+1\rangle$ with probability ϵ (i.e. completely reveals the bit a). The rest of the protocol is the same as $\mathcal{P}_{n-1}$ with the roles of Alice and Bob reversed. Thus,

$$B_n(\epsilon) \quad = \quad \epsilon + (1-\epsilon)A_{n-1}(\epsilon), \qquad \text{for } n \geq 2. \tag{8}$$

We already know from [5] that $B_1(\epsilon) = (1+\epsilon)/2$. It thus suffices to analyse Alice's cheating probability in all the protocols $\mathcal{P}_n$.

From our analysis of the three and five round protocols, we also know that

$$\begin{aligned} A_1(\epsilon) &= 1 - \frac{\epsilon}{2} \\ B_2(\epsilon) &= 1 - \frac{\epsilon}{2} + \frac{\epsilon^2}{2} \\ A_2(\epsilon) &= \frac{1}{2} + \frac{\epsilon}{2} - \frac{\epsilon^2}{2}. \end{aligned}$$

We now give a lower bound for the probability $A_n(\epsilon)$ by describing a cheating strategy for Alice that generalises the strategy that we saw in the five round protocol.

First, we assume $n = 2k-1$ ($k \geq 1$) is odd, so that Alice has k rounds of commitment, and Bob has $k-1$. The case of n even is addressed later. In the first round, Alice sends one half of the state

$$\sqrt{1-\lambda_1}\,|00\rangle \;+\; \sqrt{\frac{\lambda_1}{2}}\,|11\rangle \;+\; \sqrt{\frac{\lambda_1}{2}}\,|22\rangle$$

to Bob, and retains the other half. The half she keeps is referred to as an "entangled qutrit" below. All the parameters λ_j will be specified later. If one of the entangled qutrits she retains from any previous round is in state $x+1$, in all the commitment rounds that follow, she sends the right half of the state

$$\frac{1}{\sqrt{2}}\,|0\rangle|\psi(x,0)\rangle \;+\; \frac{1}{\sqrt{2}}\,|1\rangle|\psi(x,1)\rangle$$

and keeps the first qubit (called a "sign qubit" below). If the entangled qutrits are all 0, and at least one of the qutrits she received in earlier rounds from Bob is in state $b+1$, then she sends the right half of the state

$$\frac{1}{\sqrt{2}}\,|0\rangle|\psi(b,0)\rangle \;+\; \frac{1}{\sqrt{2}}\,|1\rangle|\psi(b,1)\rangle$$

to Bob. Otherwise, if none of the above two events occurs, she sends one half of the state

$$\sqrt{1-\lambda_j}\,|00\rangle \;+\; \sqrt{\frac{\lambda_j}{2}}\,|11\rangle \;+\; \sqrt{\frac{\lambda_j}{2}}\,|22\rangle$$

in the j-th commitment round. It is important that she do all this "in superposition," i.e., via unitary operations controlled by her entangled qutrits and the qutrits she receives from Bob. She performs no measurements in the process.

Formally, for each qutrit that Alice is supposed to send, she has a qutrit-qubit pair. The first serves as the "entangled qutrit", and the second serves as the "sign qubit". They are all initialised to the 0 state. She prepares appropriate states over these according to the above rules, as the protocol proceeds. The joint state of both parties together after the n commitment rounds then looks as given below, for an arbitrary choice of b and the signs $\{s_{2j}\}$ picked by Bob. Here, the qutrits sent by Alice to Bob are underlined. The entangled qutrits and sign bits can be identified from the context. The first two lines correspond to the part of the state where Alice commits to a bit x by sending $|x+1\rangle$, before she can identify which bit b Bob has picked. The next two lines have the part of the state where Alice has not committed to any bit, and sees a $|b+1\rangle$ in one of the qutrits Bob sent. The last term is the remaining part of the state. Note that Alice can differentiate between these three parts by examining her entangled qutrits and the qutrits Bob sent her. The odd lines contain the portion of the state constructed by Alice, and the even lines contain the portion prepared by Bob (and sent to Alice).

$$
\begin{aligned}
&\sum_{x=0,1}\sum_{j=1}^{k-1}\left[\bigotimes_{l=1}^{j-1}\sqrt{1-\lambda_l}\,|0\underline{0}\rangle\right]\otimes\sqrt{\frac{\lambda_j}{2}}\,|x+1,\underline{x+1}\rangle\otimes\left[\frac{1}{\sqrt{2}}(|0\rangle|\underline{\psi(x,0)}\rangle+|1\rangle|\underline{\psi(x,1)}\rangle)\right]^{k-j}\\
&\qquad\bigotimes\left(\sqrt{1-\epsilon}\,|0\rangle\right)^{j-1}\otimes\left[\bigotimes_{l=j}^{k-1}|\psi(b,s_{2l})\rangle\right]\\
+&\sum_{j=1}^{k-1}\left[\bigotimes_{l=1}^{j}\sqrt{1-\lambda_l}\,|0\underline{0}\rangle\right]\otimes\left[\frac{1}{\sqrt{2}}\,(|0\rangle|\underline{\psi(b,0)}\rangle+|1\rangle|\underline{\psi(b,1)}\rangle)\right]^{k-j}\\
&\qquad\bigotimes\left(\sqrt{1-\epsilon}\,|0\rangle\right)^{j-1}\otimes\left[(-1)^{s_{2j}}\sqrt{\epsilon}\,|b+1\rangle\right]\otimes\left[\bigotimes_{l=j+1}^{k-1}|\psi(b,s_{2l})\rangle\right]\\
+&\left[\bigotimes_{l=1}^{k-1}\sqrt{1-\lambda_l}\,|0\underline{0}\rangle\right]\otimes\left(\sqrt{1-\lambda_k}\,|0\underline{0}\rangle+\sqrt{\frac{\lambda_k}{2}}\,|1\underline{1}\rangle+\sqrt{\frac{\lambda_k}{2}}\,|2\underline{2}\rangle\right)\\
&\qquad\bigotimes\left(\sqrt{1-\epsilon}\,|0\rangle\right)^{k-1}.
\end{aligned}
\tag{9}
$$

Since $n=2k-1$ is odd, Bob reveals his bit b first. W.l.o.g., we may assume that Alice would like to bias the coin towards 0. She therefore sends $a=b$ in the part of her state where her entangled qutrit is not equal to $\bar{b}+1$ (which corresponds to a commitment which she cannot change). We will consider the residual state after Alice has sent back $k-i$ signs, in reverse order. This is the *unnormalised* part of the state (9) that has not been rejected by Bob. We will prove by induction that Alice can locally transform the residual state to a state $|\phi_i\rangle$ after every two rounds of sign exchange. This state is similar in form to the joint state (9) above, except that the first part is projected onto the space where $x=b$, and there is a factor of μ_i in the last term. The state $|\phi_i\rangle$ is displayed below:

$$
\begin{aligned}
&\sum_{j=1}^{i-1}\left[\bigotimes_{l=1}^{j-1}\sqrt{1-\lambda_l}\,|0\underline{0}\rangle\right]\otimes\sqrt{\frac{\lambda_j}{2}}\,|b+1,\underline{b+1}\rangle\otimes\left[\frac{1}{\sqrt{2}}(|0\rangle|\underline{\psi(b,0)}\rangle+|1\rangle|\underline{\psi(b,1)}\rangle)\right]^{i-j}\\
&\qquad\bigotimes\left(\sqrt{1-\epsilon}\,|0\rangle\right)^{j-1}\otimes\left[\bigotimes_{l=j}^{i-1}|\psi(b,s_{2l})\rangle\right]\\
+&\sum_{j=1}^{i-1}\left[\bigotimes_{l=1}^{j}\sqrt{1-\lambda_l}\,|0\underline{0}\rangle\right]\otimes\left[\frac{1}{\sqrt{2}}\,(|0\rangle|\underline{\psi(b,0)}\rangle+|1\rangle|\underline{\psi(b,1)}\rangle)\right]^{i-j}
\end{aligned}
\tag{10}
$$

$$\bigotimes \left(\sqrt{1-\epsilon}\,|0\rangle\right)^{j-1} \otimes \left[(-1)^{s_{2j}}\sqrt{\epsilon}\,|b+1\rangle\right] \otimes \left[\bigotimes_{l=j+1}^{i-1} |\psi(b,s_{2l})\rangle\right]$$

$$+ \left[\bigotimes_{l=1}^{i-1} \sqrt{1-\lambda_l}\,|0\underline{0}\rangle\right] \otimes \left(\sqrt{\mu_i(1-\lambda_i)}\,|0\underline{0}\rangle + \sqrt{\frac{\lambda_i}{2}}\,|b+1,\underline{b+1}\rangle\right)$$

$$\bigotimes \left(\sqrt{1-\epsilon}\,|0\rangle\right)^{i-1},$$

where the numbers μ_i are as follows:

$$\begin{aligned} \mu_k &= 1 \\ \mu_{i-1} &= (1-\epsilon)^2\mu_i + \frac{\epsilon}{2}(3-\epsilon). \end{aligned} \tag{11}$$

We can now also specify the parameters λ_i:

$$\lambda_i = \frac{\epsilon/2}{\mu_i(1-\epsilon)+\epsilon/2}. \tag{12}$$

Clearly, the state $|\phi_k\rangle$, when none of the signs have been revealed by Alice is of this form, with $\mu_k = 1$. Assume that this is also the case for some $i \leq k$. We will show by induction that the state after Alice has revealed $k-i+1$ sign bits $s_{2k-1}, s_{2k-3}, \ldots, s_{2i-1}$ may be transformed to (10) and that equation (11) holds.

To send the sign s_{2i-1}, Alice does the following. The sign in part of the state in the first two summations of the state $|\phi_i\rangle$ in equation (10)) is "pre-computed" (in the sign qubit). To compute the sign in the last term, Alice first sets $b+1$ to 1 in the i-th entangled qutrit, does a Hadamard transform, and exchanges that entangled qubit with the i-th sign qubit. She then measures the i-th sign qubit, and sends it across. It is easily seen that the unnormalised state that remains after Bob has checked the i-th qutrit sent by Alice is as follows.[2] Here, we have written the last terms of the first two summations in equation (10) separately in lines 5 and 7 to facilitate the rest of the proof.

$$\sum_{j=1}^{i-2} \left[\bigotimes_{l=1}^{j-1} \sqrt{1-\lambda_l}\,|0\underline{0}\rangle\right] \otimes \sqrt{\frac{\lambda_j}{2}}\,|b+1,\underline{b+1}\rangle \otimes \left[\frac{1}{\sqrt{2}}(|0\rangle|\underline{\psi(b,0)}\rangle + |1\rangle|\underline{\psi(b,1)}\rangle)\right]^{i-1-j}$$

$$\bigotimes \left(\sqrt{1-\epsilon}\,|0\rangle\right)^{j-1} \otimes \left[\bigotimes_{l=j}^{i-1} |\psi(b,s_{2l})\rangle\right]$$

$$+ \sum_{j=1}^{i-2} \left[\bigotimes_{l=1}^{j} \sqrt{1-\lambda_l}\,|0\underline{0}\rangle\right] \otimes \left[\frac{1}{\sqrt{2}}\,(|0\rangle|\underline{\psi(b,0)}\rangle + |1\rangle|\underline{\psi(b,1)}\rangle)\right]^{i-1-j}$$

$$\bigotimes \left(\sqrt{1-\epsilon}\,|0\rangle\right)^{j-1} \otimes \left[(-1)^{s_{2j}}\sqrt{\epsilon}\,|b+1\rangle\right] \otimes \left[\bigotimes_{l=j+1}^{i-1} |\psi(b,s_{2l})\rangle\right]$$

$$+ \left[\bigotimes_{l=1}^{i-2} \sqrt{1-\lambda_l}\,|0\underline{0}\rangle\right] \otimes \left(\sqrt{\frac{\lambda_{i-1}}{2}}\,|b+1,\underline{b+1}\rangle\right)$$

$$\bigotimes \left(\sqrt{1-\epsilon}\,|0\rangle\right)^{i-2} \otimes |\psi(b,s_{2(i-1)})\rangle$$

[2]Actually, the state is a mixture of two states which are both $1/\sqrt{2}$ times the state given. The mixture arises because of the two possible values of the sign bit Alice sends for s_{2i-1}. The mixture is of course equivalent to the single state shown.

$$
\begin{aligned}
&+ \left[\bigotimes_{l=1}^{i-2} \sqrt{1-\lambda_l}\,|0\underline{0}\rangle\right] \otimes \left(\sqrt{1-\lambda_{i-1}}\,|0\underline{0}\rangle\right) \\
&\qquad \bigotimes \left(\sqrt{1-\epsilon}\,|0\rangle\right)^{i-2} \otimes \left[(-1)^{s_{2(i-1)}}\sqrt{\epsilon}\,|b+1\rangle\right] \\
&+ \left[\bigotimes_{l=1}^{i-2} \sqrt{1-\lambda_l}\,|0\underline{0}\rangle\right] \otimes \left(\sqrt{1-\lambda_{i-1}}\,|0\underline{0}\rangle\right) \\
&\qquad \bigotimes \left(\sqrt{1-\epsilon}\,|0\rangle\right)^{i-2} \otimes \sqrt{1-\epsilon}\,|0\rangle(\mu_i(1-\epsilon)+\epsilon/2)^{1/2}
\end{aligned}
$$

Now, when Bob sends the sign $s_{2(i-1)}$ to Alice, she rotates the $(i-1)$-th qutrit that Bob sent her in all but the last two terms in the sum above, to $|0\rangle$. She also rotates that qutrit in the last two terms from

$$
(\mu_i(1-\epsilon)+\epsilon/2)^{1/2}\sqrt{1-\epsilon}\,|0\rangle \;+\; (-1)^{s_{2(i-1)}}\sqrt{\epsilon}\,|b+1\rangle
$$

to $\sqrt{\mu_{i-1}}|0\rangle$, whereby $\mu_{i-1} = \mu_i(1-\epsilon)^2 + \epsilon(1-\epsilon)/2 + \epsilon$. This proves the induction step.

At the final round of the protocol $\mathcal{P}_n$, the state that they are left with is

$$
\sqrt{\mu_1(1-\lambda_1)}\,|0\underline{0}\rangle \;+\; \sqrt{\frac{\lambda_1}{2}}\,|b+1,\underline{b+1}\rangle.
$$

Following Alice's strategy for computing the sign as above, we see that the probability with which Alice succeeds in passing Bob's checks is (using equation (12))

$$
\left(\sqrt{\mu_1(1-\lambda_1)(1-\epsilon)} + \sqrt{\lambda_1\epsilon/2}\right)^2 \;=\; \mu_1(1-\epsilon)+\epsilon/2. \tag{13}
$$

Solving the recurrence for μ_i given in equation (11), we get

$$
\mu_i \;=\; (1-\epsilon)^{2(k-i)} + \frac{(3-\epsilon)}{2(2-\epsilon)}(1-(1-\epsilon)^{2(k-i)}),
$$

and so that from equation (13),

$$
\begin{aligned}
\text{For } n \text{ odd,}\quad A_n(\epsilon) \;&\geq\; \frac{\epsilon}{2} + (1-\epsilon)^n + \frac{(3-\epsilon)(1-\epsilon)}{2(2-\epsilon)}(1-(1-\epsilon)^{n-1}) \\
&=\; \frac{1}{2(2-\epsilon)}(3-2\epsilon+(1-\epsilon)^{n+1}).
\end{aligned} \tag{14}
$$

The analysis in the case that $n = 2k$ is even is similar, except for the rule Alice uses to compute the bit she sends to Bob (since she is supposed to reveal her bit before Bob reveals his bit). In this case, she sends the bit x if any of her entangled qutrits is in state $|x+1\rangle$. Else, if any of Bob's qutrits is in state $b+1$, she sends b. In the remaining case, she sends $0, 1$ with equal amplitude $1/\sqrt{2}$. This leads to a state similar to $|\phi_k\rangle$ above after Bob reveals his bit, and the last sign he used, except that here $\mu_k = (1+\epsilon)/2$. So, we get:

$$
\begin{aligned}
\text{For } n \text{ even,}\quad A_n(\epsilon) \;&\geq\; \frac{\epsilon}{2} + (1-\epsilon)^{n-1}(1+\epsilon)/2 + \frac{(3-\epsilon)(1-\epsilon)}{2(2-\epsilon)}(1-(1-\epsilon)^{n-2)}) \\
&=\; \frac{1}{2(2-\epsilon)}(3-2\epsilon-(1-\epsilon)^{n+1}).
\end{aligned} \tag{15}
$$

From equations (14), (15), and (8), we can deduce lower bounds for $B_n(\epsilon)$ as well, for $n \geq 2$. This expression matches the one for $B_1(\epsilon)$. Thus,

$$\text{For all } n, \quad B_n(\epsilon) \;\geq\; \frac{1}{2(2-\epsilon)}(3-\epsilon+(-1)^n(1-\epsilon)^{n+1}). \tag{16}$$

To determine the bias achieved, we examine the maximum of the cheating probabilities attained by Alice and Bob. Note that $A_n(\epsilon) + B_n(\epsilon) \geq 3/2$ for all n, ϵ. Thus, the bias of the protocol $\mathcal{P}_n$ is at least $3/4$ for any n and ϵ. Since we would like this bias to be as small as possible, we optimise the maximum cheating probability with respect to ϵ.

For odd n, $A_n(0) = B_n(1) = 1$, and $A_n(1) = B(0) = 1/2$, A_n is monotonically decreasing, and B_n is monotonically increasing with respect to $\epsilon \in [0,1]$. Thus, the maximum bias achievable is minimised when $A_n(\epsilon) = B_n(\epsilon)$. This condition is satisfied when $\epsilon = \epsilon_0$ such that

$$\begin{aligned} (1-\epsilon_0)^{n+1} &= \frac{\epsilon_0}{2} \qquad \text{that is,} \\ \epsilon_0 &= \frac{1}{n+1}(\ln n - \ln\ln n + \Theta(1)). \end{aligned} \tag{17}$$

Using equation (17) we can verify that the lower bound on $A_n(\epsilon_0)$ is $3/4$.

For even n, $A_n(0) = A_n(1) = 1/2$, and $B(0) = B_n(1) = 1$, A_n is concave, B_n is convex, and $B_n(\epsilon) \geq A_n(\epsilon)$ for $\epsilon \in [0,1]$. Thus, the bias achievable is minimised when B_n is, i.e., for $\epsilon = \epsilon_0$ such that

$$\begin{aligned} (1-\epsilon_0)^{n} &= \frac{1}{(2-\epsilon_0)n+1} \qquad \text{that is,} \\ \epsilon_0 &= \frac{1}{n}\ln(2n - \ln n + \Theta(1)). \end{aligned} \tag{18}$$

The expression for the cheating probability $B_n(\epsilon_0)$ then evaluates to at least $\frac{3}{4}$, as may be seen by using equation (18).

This completes the analysis of the cheating strategies devised above.

Acknowledgements

A.N. was supported by Charles Lee Powell Foundation, and NSF grants CCR 0049092 and EIA 0086038. Part of this work was done while this author was at DIMACS Center, Rutgers University, and AT&T Labs, and was supported by NSF grants STC 91-19999, CCR 99-06105 and EIA 00-80234.

References

[1] M. Blum. Coin flipping by telephone: A protocol for solving impossible problems. *Advances in Cryptology: Report on CRYPTO'81*, pp. 11–15.

[2] H. Lo and H. Chau. Why quantum bit commitment and ideal quantum coin tossing are impossible. *Physica D*, 120:177–187, 1998. See also quant-ph/9711065.

[3] D. Mayers. Unconditionally secure quantum bit commitment is impossible. *Physical Review Letters*, 78:3414–3417, 1997.

[4] D. Aharonov, A. Ta-Shma, U. Vazirani, and A. Yao. In *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing*, pp. 705–714, 2000.

[5] A. Ambainis. A new protocol and lower bounds for quantum coin flipping. In *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing*, pp. 134–142, 2001.

[6] A.C.-C. Yao. Quantum circuit complexity. In *Proceedings of the 34th Annual Symposium on Foundations of Computer Science*, pp. 352–361, 1993.

[7] L. Salvail. Communicated by Andris Ambainis, 2001.

[8] E. Bernstein and U. Vazirani. Quantum complexity theory. *SIAM Journal on Computing*, 26(5):1411–1473, October 1997.

[9] D. Mayers, L. Salvail, and Y. Chiba-Kohno. Unconditionally secure quantum coin-tossing. LANL Preprint quant-ph/9904078.

[10] A. Kitaev. Personal communication, 2001.

[11] I. Kerenidis and A. Nayak. Manuscript, 2001.

[12] A. Ambainis. Personal communication, 2001.

[13] R.W. Spekkens and T. Rudolph. A quantum protocol for cheat-sensitive weak coin flipping. LANL Preprint quant-ph/0202118.

[14] R. Jozsa. Fidelity for mixed quantum states. *Journal of Modern Optics*, 41(12):2315–2323, 1994.

[15] C.A. Fuchs and J. van de Graaf. Cryptographic distinguishability measures for quantum-mechanical states. *IEEE Transactions on Information Theory*, 45(4):1216–1227, 1999.