Notes on canceling coherent errors

IM

(Dated: February 17, 2019)

Let U_θ be the unitary that describes the coherent error on the system, where θ is a random variable with probability density $p(\theta)$. We assume the density $p(\theta)$ and the dependence of U_θ to θ are known, but the actual value of θ is unknown. To obtain information about θ we perform measurements on the *spectator* qubits. Suppose these measurements yield outcome $m \in \mathcal{M}$, with probability $p(m|\theta)$, where $\mathcal{M}$ is a finite set of possible outcomes (For instance, it could be a finite bit string corresponding to the outcomes of the measurements on spectator qubits). The conditional distribution $p(m|\theta)$ is determined by our choice of state preparation and measurement on the spectator qubits.

Then, for each possible outcome m, we choose an appropriate unitary V_m, which approximately cancels the effect of U_θ, such that $V_m U_\theta$ is approximately equal to the identity operator for most values of θ and m. To summarize, our protocol for correcting coherent errors is determined by a conditional probability $p(m|\theta)$ and the choice of unitaries $\{V_m\}$.

To quantify the performance of a protocol we can use the average gate fidelity or, equivalently, entanglement fidelity. Recall that the entanglement fidelity for a channel $\mathcal{E}$ with Kraus operators $\{E_k\}$, is given by

$$F_e(\mathcal{E}) = \langle\Psi|[\mathcal{E}\otimes\mathcal{I}](|\Psi\rangle\langle\Psi|)|\Psi\rangle = \frac{1}{d^2}\sum_k |\mathrm{Tr}(E_k)|^2 \ , \tag{1}$$

where $|\Psi\rangle$ is a maximally-entangled state, and d is the dimension of the system Hilbert space. This is related to the average gate fidelity, via the formula

$$\int d\psi \ \langle\psi|\mathcal{E}(|\psi\rangle\langle\psi|)|\psi\rangle = \frac{d+\sum_k |\mathrm{Tr}(E_k)|^2}{d(d+1)} = \frac{1+d\times F_e(\mathcal{E})}{d+1} \ , \tag{2}$$

where $d\psi$ is the uniform (unitarily invariant) measure over pure states ([1], [2] and [3].)

To quantify the overall error after applying the unitary correction V_m, we consider the average entanglement fidelity. Assuming the conditional probability $p(m|\theta)$ is given and fixed, we write this average as

$$\overline{F}(\{V_m\}) = \int d\theta\, p(\theta) \sum_{m\in\mathcal{M}} p(m|\theta)\times F_e(V_m U_\theta) \tag{3}$$

$$= \int d\theta \sum_{m\in\mathcal{M}} p(\theta,m)\times \frac{1}{d^2}|\mathrm{Tr}(V_m U_\theta)|^2 \ , \tag{4}$$

where $F_e(V_m U_\theta)$ denotes the entanglement fidelity for the unitary transformation $V_m U_\theta$, and $p(\theta,m) = p(m|\theta)p(\theta)$.

The fact that there are a finite number of spectator qubits, limits the form of the conditional distribution $p(m|\theta)$ that can be achieved by performing measurements on these qubits. Our final goal is to optimize over all such realizable conditional distributions $p(m|\theta)$ and unitaries $\{V_m\}$ to find the maximum achievable value of average entanglement fidelity

$$\max_{p(m|\theta)}\max_{\{V_m\}} \ \overline{F}(\{V_m\}) = \max_{p(m|\theta)}\max_{\{V_m\}} \ \int d\theta\, p(\theta) \sum_{m\in\mathcal{M}} p(m|\theta)\times\frac{1}{d^2}|\mathrm{Tr}(V_m U_\theta)|^2 \ , \tag{5}$$

This optimization problem can be solved in two steps: first, we assume $p(m|\theta)$ is fixed and given, and find the set $\{V_m\}$ which maximizes the average gate fidelity for this $p(m|\theta)$, and then we maximize over the set of all possible conditionals $p(m|\theta)$. In the first step of the optimization problem, where $p(m|\theta)$ is fixed, we need to solve

$$\overline{F}_{\mathrm{opt}} = \max_{\{V_m\}} \ \int d\theta\, p(\theta) \sum_{m\in\mathcal{M}} p(m|\theta)\times\frac{1}{d^2}|\mathrm{Tr}(V_m U_\theta)|^2 \ , \tag{6}$$
$$\text{subject to}: \ V_m V_m^\dagger = I, \ \forall m\in\mathcal{M} \ .$$

As we will discuss later, in the case of a single qubit system, this optimization problem can be easily solved. We also consider the general case of $d > 2$, and study this constrained optimization problem.

It is useful to note that we can optimize over unitaries V_m independently, i.e. for each m we can find the optimal

$$\max_{V_m} \int d\theta\, p(\theta)p(m|\theta) \times \frac{1}{d^2}|\mathrm{Tr}(V_m U_\theta)|^2 \ , \tag{7}$$

$$\text{subject to}: \ V_m V_m^\dagger = I \ , \tag{8}$$

and the set of these unitaries will form the optimal set $\{V_m\}$ in Eq.(5). This optimization problem has a simple interpretation: Assuming we have observed outcome m on the spectator qubits, the noise on the main system can be describe by the quantum channel

$$\mathcal{E}_m(\cdot) \equiv \int d\theta\, p(\theta|m)\, U_\theta(\cdot)U_\theta^\dagger \tag{9}$$

where

$$p(\theta|m) = \frac{p(\theta)p(m|\theta)}{\int d\theta\, p(\theta)p(m|\theta)} = \frac{p(\theta)p(m|\theta)}{p(m)} \tag{10}$$

is the conditional probability of parameter θ, given that we have observed outcome $m \in \mathcal{M}$ on the spectator qubits. Then, the goal is to find a unitary V_m which approximately cancels the effect of $\mathcal{E}_m$. I.e. we want to maximize the entanglement fidelity

$$F_e(\tilde{\mathcal{E}}_m) = \int d\theta\, p(\theta|m) \times \frac{1}{d^2}|\mathrm{Tr}(V_m U_\theta)|^2 \ , \tag{11}$$

where

$$\tilde{\mathcal{E}}_m(\cdot) \equiv \mathcal{V}_m \circ \mathcal{E}_m(\cdot) \equiv V_m[\mathcal{E}_m(\cdot)]V_m^\dagger \ . \tag{12}$$

We conclude that

$$\overline{F}(\{V_m\}) = \sum_{m\in\mathcal{M}} p(m)F_e(\tilde{\mathcal{E}}_m) = \sum_{m\in\mathcal{M}} p(m)F_e(\mathcal{V}_m \circ \mathcal{E}_m) \ , \tag{13}$$

and

$$\overline{F}_{\text{opt}} = \max_{\{V_m\}} \sum_{m\in\mathcal{M}} p(m)F_e(\mathcal{V}_m \circ \mathcal{E}_m) = \sum_{m\in\mathcal{M}} p(m) \max_{V_m} F_e(\mathcal{V}_m \circ \mathcal{E}_m), \tag{14}$$

$$\text{subject to}: \ V_m V_m^\dagger = I, \ \forall m \in \mathcal{M} \ ,$$

where $\mathcal{V}_m(\cdot) = V_m(\cdot)V_m^\dagger$

Using this interpretation, we can easily find the optimal unitaries $\{V_m\}$ in the qubit case.

THE OPTIMAL STRATEGY FOR THE QUBIT CASE

In the following, we find the optimal unitaries $\{V_m\}$ for a given conditional probability $p(m|\theta)$, in the case of a qubit ($d = 2$).

Liouville representation of quantum channels

We use the Liouville representation of a quantum channel in the Pauli basis. For a qubit system, normally the Liouville representation is the 4×4 matrix which determines the action of the channel on the Pauli operators and the identity operator. However, because all the relevant channels in our discussion are unital, i.e. they map the identity operator to itself, we can focus on the 3×3 sub-matrices, which determine the action of channels $\mathcal{E}_m$ on the Pauli operators.

Let $\Lambda(\mathcal{E})$ be the the 3×3 matrix of the Liouville representation of the channel $\mathcal{E}$, defined by

$$\Lambda(\mathcal{E})_{a,b} = \frac{1}{2}\mathrm{Tr}(\sigma_a \mathcal{E}(\sigma_b)) \ , \tag{15}$$

where $a, b \in 1, 2, 3$. The normalization is chosen such that for the identity channel the Liouville representation is the identity matrix. It can be easily seen that for a unital channel $\mathcal{E}$, it holds that

$$F_e(\mathcal{E}) = \frac{1 + \text{Tr}(\Lambda(\mathcal{E}))}{4} . \tag{16}$$

To see this, consider the maximally-entangled state $|\Psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$. Then,

$$|\Psi^-\rangle\langle\Psi^-| = \frac{1}{4}(I \otimes I - [\sigma_1 \otimes \sigma_1 + \sigma_2 \otimes \sigma_2 + \sigma_3 \otimes \sigma_3]) , \tag{17}$$

Then, using the orthogonality of Pauli operators $\text{Tr}(\sigma_a \sigma_b) = 2\delta_{a,b}$, we obtain

$$F_e(\mathcal{E}) = \langle\Psi^-|[\mathcal{E} \otimes \mathcal{I}](|\Psi^-\rangle\langle\Psi^-|)|\Psi^-\rangle = \frac{1}{16}(4 + 4\,\text{Tr}(\Lambda(\mathcal{E}))) = \frac{1}{4}(1 + \text{Tr}(\Lambda(\mathcal{E}))) , \tag{18}$$

which proves Eq.(16) [1].

Our goal is to find the unitary V, which maximizes the entanglement fidelity of unital channel $\mathcal{E}$,

$$\max_V F_e(\mathcal{V} \circ \mathcal{E}) , \tag{19}$$

where $\mathcal{V}(\cdot) = V(\cdot)V^\dagger$. In the following we use some useful properties of the Liouville representation to answer this question: (i) Liouville representation in the Pauli basis is a real matrix. (ii) If unital channel $\mathcal{E}_1$ has representation $\Lambda(\mathcal{E}_1)$ and the unital channel $\mathcal{E}_2$ has representation $\Lambda(\mathcal{E}_2)$, then the channel $\mathcal{E}_2 \circ \mathcal{E}_1$ has representation $\Lambda(\mathcal{E}_2)\Lambda(\mathcal{E}_1)$, and $\mathcal{E}_2 + \mathcal{E}_1$ has representation $\Lambda(\mathcal{E}_2) + \Lambda(\mathcal{E}_1)$. (iii) For a unitary channel $\mathcal{V}(\cdot) = V(\cdot)V^\dagger$, the Liouville representation (in the Pauli basis) is an element of SO(3), i.e. $\Lambda(\mathcal{V})\Lambda(\mathcal{V})^T = I$. Conversely, for any given 3×3 matrix $O \in$ SO(3), there exists a 2×2 unitary matrix V, such that the representation of the channel $\mathcal{V}(\cdot) = V(\cdot)V^\dagger$, is O (Later, we will see how we can find the unitary V corresponding to a given $O \in$ SO(3)).

Suppose for a given unital channel $\mathcal{E}$, we are interested to find a unitary V such that $\mathcal{V} \circ \mathcal{E}$ has the largest entanglement fidelity, where $\mathcal{V}(\cdot) = V(\cdot)V^\dagger$. The Liouville representation for $\mathcal{V} \circ \mathcal{E}$ is $\Lambda(\mathcal{V})\Lambda(\mathcal{E})$. Then, the entanglement fidelity of $\mathcal{V} \circ \mathcal{E}$ is equal to

$$F_e(\mathcal{V} \circ \mathcal{E}) = \frac{1 + \text{Tr}(\Lambda(\mathcal{V})\Lambda(\mathcal{E}))}{4} . \tag{20}$$

Since V is an arbitrary qubit unitary, $\Lambda(\mathcal{V})$ can be an arbitrary element of SO(3). Then, to maximize $F_e(\mathcal{V} \circ \mathcal{E})$ we need to choose $O = \Lambda(\mathcal{V})$ to be an element of SO(3) for which $\text{Tr}(O\,\Lambda(\mathcal{E}))$ is maximized.

To find the optimal O, we use the following fact: For any given $n \times n$ matrix A, and unitary S,

$$|\text{Tr}(SA)| \leq \|A\|_1 = \text{Tr}(\sqrt{A^\dagger A}) , \tag{21}$$

where $\|\cdot\|_1$ denotes the l_1-norm. Furthermore, the equality can be achieved for $S = \frac{1}{\sqrt{A^\dagger A}} A^\dagger$ (The fact that $\frac{1}{\sqrt{A^\dagger A}} A^\dagger$ is unitary can be shown, e.g., using the Singular Value Decomposition of A). Note that here we are assuming matrix A is invertible. In the context of the above discussion, this assumption is justified, because otherwise the channel is too noisy (entanglement-breaking). If matrix A is real, then $A^\dagger = A^T$, where T denotes the transpose. In this case, the unitary matrix $S = \frac{1}{\sqrt{A^T A}} A^T$ is also real, and hence is an element of O(n). Assuming $\det(A) > 0$, we find that S has determinant $+1$, and therefore is an element of SO(n).

Therefore, using the fact that $\Lambda(\mathcal{E})$ is a real matrix, we can show that the maximum entanglement fidelity is equal to

$$\max_{O \in \text{SO(3)}} \text{Tr}(O\,\Lambda(\mathcal{E})) = \|\Lambda(\mathcal{E})\|_1 = \text{Tr}(\sqrt{\Lambda(\mathcal{E})^T \Lambda(\mathcal{E})}) , \tag{22}$$

Furthermore, the rotation $O \in$ SO(3), for which this maximum is achieved can be chosen to be

$$O_{\mathcal{E}} = \frac{1}{\sqrt{\Lambda(\mathcal{E})^T \Lambda(\mathcal{E})}} \Lambda(\mathcal{E})^T . \tag{23}$$

[1] Note that for non-unital channels a generalization of Eq.(16) remains valid for the 4×4 Liouville representation $\tilde{\Lambda}(\mathcal{E})$ in the Pauli basis. Then, $F_e(\mathcal{E}) = \text{Tr}(\Lambda(\mathcal{E}))/4$. This can also be easily generalized to the case of multi-qubit systems.

Here, we have assumed $\sqrt{\Lambda(\mathcal{E})^T\Lambda(\mathcal{E})}$ is invertible, which is justified because otherwise the channel $\mathcal{E}$ will be too noisy, and in fact, it will be entanglement-breaking.

Next, we determine the 2×2 unitary $V\in$ SU(2) whose corresponding Liouville representation is equal to $O_{\mathcal{E}}$. To find the unitary V, we can use the Euler decomposition of $O_{\mathcal{E}}$, or we can use the following approach based on the eigen-decomposition of $O_{\mathcal{E}}$: As an element of SO(3), matrix $O_{\mathcal{E}}$ has an eigenvalue 1, and two eigenvalues $e^{\pm i\phi}$ with $0\leq\phi\leq\pi$ (Because $O_{\mathcal{E}}$ is a real matrix, its eigenvalues come in pair, and because it is an element of SO(3), it should have determinant 1, which implies the product of its eigenvalues is 1. This means that the eigenvalues should be $1, e^{i\pm\phi}$). Furthermore, the eigenvector corresponding to eigenvalue 1, can be chosen to have only real elements. Let $\hat{n}=(n_1,n_2,n_3)$ be the normalized eigenvector of $O_{\mathcal{E}}$ corresponding to eigenvalues 1, with $n_1,n_2,n_3\in\mathbb{R}$. Then, the 2×2 unitary V corresponding to $O_{\mathcal{E}}$ can be chosen to be one of the unitaries

$$O_{\mathcal{E}}=\frac{1}{\sqrt{\Lambda(\mathcal{E})^T\Lambda(\mathcal{E})}}\Lambda(\mathcal{E})^T \quad\longrightarrow\quad V=e^{\pm i\frac{\phi}{2}(n_1\sigma_1+n_2\sigma_2+n_3\sigma_3)}=\cos\frac{\phi}{2}I\pm i\sin\frac{\phi}{2}[n_1\sigma_1+n_2\sigma_2+n_3\sigma_3]\ . \tag{24}$$

Note that the factor $1/2$ in the exponent is related to the fact that V can be thought as the spin-half representation of rotation $O_{\mathcal{E}}\in$ SO(3).

To determine which one of these two unitaries should be chosen, we can use the following method: To simplify the discussion assume $n_3\neq 0$ (otherwise, we can repeat the argument with another component of the eigenvector $\hat{n}$). Note that if (n_1,n_2,n_3) is an eigenvector of $O_{\mathcal{E}}$, then $(-n_1,-n_2,-n_3)$ is also an eigenvector. Using this freedom, we can choose the eigenvector such that $n_3>0$. Then, if the matrix element $[O_{\mathcal{E}}]_{2,1}$ is positive, we choose $V=e^{-i\frac{\phi}{2}(n_1\sigma_1+n_2\sigma_2+n_3\sigma_3)}$, and otherwise we choose $V=e^{i\frac{\phi}{2}(n_1\sigma_1+n_2\sigma_2+n_3\sigma_3)}$.

To summarize, given a unital qubit channel $\mathcal{E}$, the qubit unitary channel $\mathcal{V}(\cdot)\equiv V(\cdot)V^\dagger$, which maximizes the entanglement fidelity $F_e(\mathcal{V}\circ\mathcal{E})$, is determined by Eq.(24), where the SO(3) matrix $O_{\mathcal{E}}=\frac{1}{\sqrt{\Lambda(\mathcal{E})^T\Lambda(\mathcal{E})}}\Lambda(\mathcal{E})^T$ is obtained from the Liouville representation $\Lambda(\mathcal{E})$ associated to $\mathcal{E}$, defined in Eq.(15). Furthermore, the entanglement fidelity $F_e(\mathcal{V}\circ\mathcal{E})$ for this optimal choice of V, is equal to

$$F_e(\mathcal{V}\circ\mathcal{E})=\frac{1+\|\Lambda(\mathcal{E})\|_1}{4}\ . \tag{25}$$

The matrix $O_{\mathcal{E}}$ has a simple geometric interpretation. Let $\vec{r}\in\mathbb{R}^3$ be a vector in the Bloch sphere. Then, the action of a general qubit channel on the Bloch sphere is $\vec{r}\rightarrow SM\vec{r}+\vec{c}$, where M is a positive real matrix which contracts the vector $\vec{r}$, $S\in$ SO(3) is a rotation, and $\vec{c}\in\mathbb{R}^3$ is a shift (See [3]). For the unital channels, $\vec{c}=0$. By applying a unitary rotation we can cancel the rotation S. The argument presented above shows that this is the optimal strategy, which maximizes the entanglement fidelity, i.e. we cannot compensate for the effect of the contraction M by applying a unitary.

The above method determines the unitary V_m which maximizes the entanglement fidelity in Eq.(13). This leads to the following result:

Result: Let $\Lambda_\theta\in$ SO(3), defined by $[\Lambda_\theta]_{a,b}=\frac{1}{2}\mathrm{Tr}(\sigma_a U_\theta\sigma_b U_\theta^\dagger)$, be 3×3 Liouville matrix associated to the random noise unitary U_θ. Then, the maximum average fidelity after applying the correcting unitaries $\{V_m\}$ (defined in Eq.14) is given by

$$\overline{F}_{\text{opt}}=\frac{1+\sum_m\left\|\int d\theta\ p(\theta)p(m|\theta)\Lambda_\theta\right\|_1}{4}\ . \tag{26}$$

On the other hand, without applying the correcting unitaties $\{V_m\}$, the entanglement fidelity is

$$F=\frac{1+\int d\theta\ p(\theta)\mathrm{Tr}(\Lambda_\theta)}{4}\ , \tag{27}$$

Using this result, finding the optimal strategy reduces to finding the state-preparation and measurement on the spectator qubits, which maximizes $\sum_m\left\|\int d\theta\ p(\theta)p(m|\theta)\Lambda_\theta\right\|_1$.

Special case of random rotations around a fixed axis

Next, we consider the special case where the noise is a random rotation around a fixed, but arbitrary axis. Assume the noise unitary is in the form of

$$U_\theta=e^{i\sigma_3 f(\theta)}\ , \tag{28}$$

i.e. a rotation around z, with angle $f(\theta)$, which occurs with probability density $p(\theta)$. Then, for $U_\theta = e^{if(\theta)\sigma_z}$, the corresponding Liouville operator is

$$\Lambda_\theta = \Lambda(U_\theta) = \begin{pmatrix} \cos 2f(\theta) & -\sin 2f(\theta) & \\ \sin 2f(\theta) & \cos 2f(\theta) & \\ & & 1 \end{pmatrix} \tag{29}$$

Suppose the information obtained from the spectator qubits is described by the conditional distribution $p(m|\theta)$. Then, the operator $\int d\theta\; p(\theta)p(m|\theta)\Lambda_\theta$ can be written as

$$\int d\theta\; p(\theta)p(m|\theta)\;\Lambda_\theta = \int d\theta\; p(\theta)p(m|\theta) \begin{pmatrix} \cos 2f(\theta) & -\sin 2f(\theta) & \\ \sin 2f(\theta) & \cos 2f(\theta) & \\ & & 1 \end{pmatrix} . \tag{30}$$

This operator can be easily diagonalized. Its eigenvalues are

$$\Big\{ p(m) = \int d\theta\; p(\theta)p(m|\theta) \;,\; \int d\theta\; p(\theta)p(m|\theta)\; e^{\pm i2f(\theta)} \Big\} \tag{31}$$

Therefore, using Eq.(26) we conclude that

$$\overline{F}_{\text{opt}} = \frac{1}{4} + \frac{1}{4}\Big[\sum_m \int d\theta\; p(\theta)p(m|\theta) + 2\sum_m |\int d\theta\; p(\theta)p(m|\theta)\; e^{i2f(\theta)}| \Big] \tag{32}$$

$$= \frac{1}{2} + \frac{1}{2}\sum_m \Big| \int d\theta\; p(\theta)p(m|\theta)\; e^{i2f(\theta)} \Big| \tag{33}$$

Note that without applying the correcting unitaries $\{V_m\}$, the average fidelity is

$$F = \frac{1}{2} + \frac{1}{2}\int d\theta\; p(\theta)\cos[2f(\theta)] \;. \tag{34}$$

In summary, we find that

Result: Suppose the noise unitary $e^{if(\theta)\sigma_z}$ affects the system with the probability density $p(\theta)$. Suppose the information obtained from the spectator qubits is described by the conditional distribution $p(m|\theta)$. Then, by applying a set of unitaries $\{V_m\}$ we can achieve the optimal average entanglement fidelity equal to

$$\overline{F}_{\text{opt}} = \frac{1}{2} + \frac{1}{2}\sum_m \Big| \int d\theta\; p(\theta)p(m|\theta)\; e^{i2f(\theta)} \Big| \tag{35}$$

The optimal unitaries $\{V_m\}$ are determined by $V_m = e^{i\phi_m\sigma_z}$, where

$$\phi_m = -\frac{1}{2}\arctan\frac{\int d\theta\; p(\theta)p(m|\theta)\sin[2f(\theta)]}{\int d\theta\; p(\theta)p(m|\theta)\cos[2f(\theta)]} \;. \tag{36}$$

On the other hand, without applying these unitaries the average entanglement fidelity is equal to

$$F = \frac{1}{2} + \frac{1}{2}\int d\theta\; p(\theta)\cos[2f(\theta)] \;. \tag{37}$$

In particular, applying the unitaries $\{V_m\}$ helps whenever for some $m \in \mathcal{M}$

$$\int d\theta\; p(\theta)p(m|\theta)\sin[2f(\theta)] \neq 0 \;. \tag{38}$$

Eq.(35) determines the dependence of the optimal fidelity to the conditional distribution $p(m|\theta)$. We can use this formula to determine the optimal measurement on the spectator qubits.

A simple example with one and two spectator qubits

Suppose the noise is described by the unitary operator $U_\theta = e^{i\sigma_3\theta}$, i.e. a rotation around z, with angle θ. Assume θ is uniformly distributed between $-\alpha$, and α, i.e.

$$p(\theta) = \frac{1}{2\alpha} \ : \ |\theta| \leq \alpha \, , \tag{39}$$

and zero otherwise, where $0 < \alpha < \pi$.

First, we find the original entanglement fidelity, i.e. the entanglement fidelity without applying the correcting unitaries. This is equal to

$$F = \frac{1}{2} + \frac{\sin 2\alpha}{4\alpha} \, . \tag{40}$$

Note that for $\alpha \to 0$, the fidelity goes to one, and for $\alpha = \pi$, $F = 1/2$. This corresponds to the case of full-dephasing channel.

Next, assume we have access to a single spectator qubit. For simplicity, we assume the spectator qubit is subjected to exactly the same noise. Furthermore, we assume we initially prepare the spectator qubit in state $(|0\rangle + |1\rangle)/\sqrt{2}$, and after noise $U_\theta = e^{i\theta\sigma_3}$, we measure it in the eigenbasis of σ_y. The probability of two outcomes are given by

$$p(m=0|\theta) = |\cos(\theta + \pi/4)|^2 = \frac{1}{2}(1 - \sin(2\theta)) \tag{41}$$

$$p(m=1|\theta) = 1 - |\cos(\theta + \pi/4)|^2 = \frac{1}{2}(1 + \sin(2\theta)) \, . \tag{42}$$

Then,

$$\int d\theta \; p(\theta) p(m=0|\theta) \; e^{i2\theta} = \frac{1}{2\alpha} \int_{-\alpha}^{\alpha} d\theta \; \frac{1}{2}(1 - \sin(2\theta)) e^{i2\theta} \tag{43}$$

$$= \frac{1}{4\alpha} \Big[\sin(2\alpha) - i(\alpha - \frac{\sin(4\alpha)}{4}) \Big] \tag{44}$$

and, similarly,

$$\int d\theta \; p(\theta) p(m=1|\theta) \; e^{i2\theta} = \frac{1}{4\alpha} \Big[\sin(2\alpha) + i(\alpha - \frac{\sin(4\alpha)}{4}) \Big] \tag{45}$$

Therefore,

$$\Big| \int d\theta \; p(\theta) p(m=0|\theta) \; e^{i2\theta} \Big| = \Big| \int d\theta \; p(\theta) p(m=1|\theta) \; e^{i2\theta} \Big| \tag{46}$$

$$= \frac{1}{4\alpha} [\sqrt{\sin^2(2\alpha) + [\alpha - \frac{\sin(4\alpha)}{4}]^2}] \tag{47}$$

$$= \frac{|\sin(2\alpha)|}{4\alpha} [\sqrt{1 + (\frac{\alpha - \frac{\sin(4\alpha)}{4}}{\sin(2\alpha)})^2}] \, , \tag{48}$$

which implies

$$\overline{F}_{\text{opt}} = \frac{1}{2} + \frac{1}{2} \sum_m \Big| \int d\theta \; p(\theta) p(m|\theta) \; e^{i2f(\theta)} \Big| \tag{49}$$

$$= \frac{1}{2} + \frac{\sin(2\alpha)}{4\alpha} \sqrt{1 + [\frac{4\alpha - \sin(4\alpha)}{4\sin(2\alpha)}]^2} \tag{50}$$

Then, for small values of α, we find

$$F \approx 1 - \frac{\alpha^2}{3} \tag{51}$$

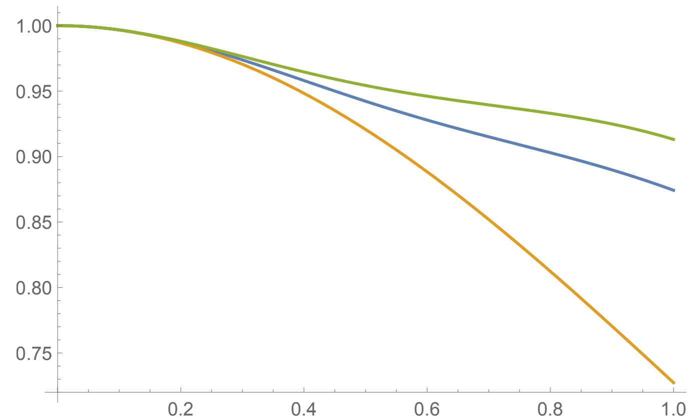

FIG. 1. Average entanglement fidelity as a function of angle α with zero, one and two spectator qubits. We assume the noise on all qubits is described by a random unitary $e^{i\theta\sigma_z}$, where θ is uniformly distributed in the interval $[-\alpha, \alpha]$, and $0 \leq \alpha < \pi$. The orange curve shows the actual average entanglement fidelity without applying any correction. The blue curve corresponds to the case where we use one spectator qubit to obtain information about the noise. We assume the spectator qubit is initially prepared in state $(|0\rangle + |1\rangle)/\sqrt{2}$ and then is measured in σ_y basis. Then based on the outcome of the measurement we apply a unitary correction to the main qubit. The green curve shows the case where we use two spectator qubits. Again we assume both spectator qubits are initially prepared in state $(|0\rangle + |1\rangle)/\sqrt{2}$ and measured in the σ_y basis. Then, based on the outcomes of these measurements we apply a unitary transformation on the main qubit to correct the rotation $e^{i\theta\sigma_z}$.

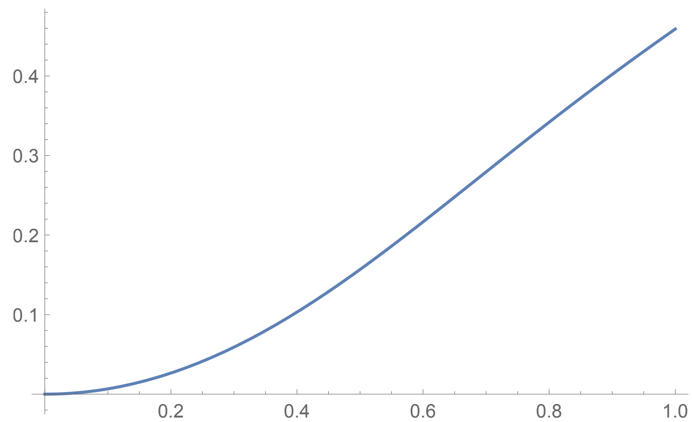

FIG. 2. The correction angle $\phi_0 = -\phi_1$ as a function of α, in the case of single spectator qubit. The horizontal axis is the angle α as defined in the above figure. The vertical axis is the angle $\phi_0 = -\phi_0$, found in Eq.(53). To correct the effect of noise we apply the unitary $e^{i\phi_0\sigma_z}$ in the case of outcome 0, and the unitary $e^{-i\phi_0\sigma_z}$ in the case of outcome 1.

and

$$\overline{F}_{\text{opt}} - F \approx \frac{4}{9} \times \alpha^4 \ . \tag{52}$$

To achieve this optimal fidelity, we need to apply the unitary $e^{i\phi_m\sigma_z}$, where $m \in 0, 1$ is the outcome of the σ_y measurement on the spectator qubit. Using Eq.(36) we find that

$$\phi_0 = -\phi_1 = -\frac{1}{2} \arctan \frac{\int_{-\alpha}^{\alpha} d\theta \ [1 + \sin(2\theta)] \sin(2\theta)}{\int_{-\alpha}^{\alpha} d\theta \ [1 + \sin(2\theta)] \cos(2\theta)} \ = \frac{1}{2} \arctan \frac{4\alpha - \sin(4\alpha)}{4\sin(2\alpha)} \ . \tag{53}$$

This angle is plotted in Fig.2. It is interesting to note that for small $\alpha \ll 1$, this angle is proportional α^2, i.e.

$$\phi_0 = -\phi_1 \approx \frac{2}{3}\alpha^2 \ , \tag{54}$$

whereas for large α it increases linearly with α.

Next, we assume we have access to two spectator qubits. Again, we assue the same unitary $e^{i\theta\sigma_z}$ is applied on both qubits. Also, we assume both qubits are initially prepared in state $(|0\rangle+|1\rangle)/\sqrt{2}$, and are measured in σ_y basis. Then, there are four outcomes with probabilities

$$p(m=00|\theta)=\frac{1}{4}[1-\sin(2\theta)]^2$$

$$p(m=01|\theta)=\frac{1}{4}[1-\sin^2(2\theta)] \tag{55}$$

$$p(m=10|\theta)=\frac{1}{4}[1-\sin^2(2\theta)] \tag{56}$$

$$p(m=11|\theta)=\frac{1}{4}[1+\sin(2\theta)]^2 \ . \tag{57}$$

Putting these conditional probabilities in Eq.(35) we can find the optimal achievable average fidelity (See green curve in Fig.(1)).

[To be added: Plot the error in terms of the diamond norm]

A GENERAL UPPER BOUND ON THE ACHIEVABLE ENTANGLEMENT FIDELITY

The above results are limited to the case of a qubit ($d=2$). Next, we consider the optimization problem in the general case. First, we derive a simple upper-bound on $\overline{F}_{\text{opt}}$.

$$\overline{F}_{\text{opt}}\leq\frac{1}{d^2}\sum_{m\in\mathcal{M}}\Big\|\int d\theta\, p(\theta,m)[U_\theta\otimes U_\theta^*]\Big\|_1\leq 1\ , \tag{58}$$

where $p(\theta,m)=p(\theta)p(m|\theta)$ [2].

The first inequality follows from the fact that for any choice of unitaries $\{V_m\}$, it holds that

$$\int d\theta\, p(\theta)\sum_{m\in\mathcal{M}}p(m|\theta)\times\frac{1}{d^2}|\text{Tr}(V_mU_\theta)|^2=\frac{1}{d^2}\sum_{m\in\mathcal{M}}\int d\theta\, p(\theta,m)\ |\text{Tr}(V_mU_\theta)|^2 \tag{59}$$

$$=\frac{1}{d^2}\sum_{m\in\mathcal{M}}\times\text{Tr}\Big(\big(\int d\theta\, p(\theta,m)[U_\theta\otimes U_\theta^*]\big)[V_m\otimes V_m^*]\Big) \tag{60}$$

$$\leq\frac{1}{d^2}\sum_{m\in\mathcal{M}}\Big\|\int d\theta\, p(\theta,m)[U_\theta\otimes U_\theta^*]\Big\|_1\ , \tag{61}$$

where to get the last line we have used the fact that for any unitary S, $|\text{Tr}(AS)|\leq\|A\|_1$. The second inequality in Eq.58 follows from the triangle inequality,

$$\frac{1}{d^2}\sum_{m\in\mathcal{M}}\|\int d\theta\, p(\theta,m)[U_\theta\otimes U_\theta^*]\|_1\leq\frac{1}{d^2}\sum_{m\in\mathcal{M}}\int d\theta\, p(\theta,m)\big\|[U_\theta\otimes U_\theta^*]\big\|_1 \tag{62}$$

$$\leq\frac{1}{d^2}\sum_{m\in\mathcal{M}}\int d\theta\, p(\theta,m)\times d^2 \tag{63}$$

$$\leq 1\ , \tag{64}$$

where we have used the fact that $\big\|[U_\theta\otimes U_\theta^*]\big\|_1=\|U_\theta\|_1^2=d^2$. This completes the proof of Eq.58.

---

[2] Using the same argument, we can also prove

$$F_{\text{opt}}\leq\frac{1}{d^2}\sum_{m\in\mathcal{M}}\|\int d\theta\, p(\theta,m)[U_\theta\otimes U^\dagger(\theta)]\big\|_1\leq 1\ ,$$

[I need to check if these two conditions are really different or not]

A GENERAL (SUB-OPTIMAL) STRATEGY FOR CORRECTING COHERENT ERRORS

In the case of a qubit, we have found the optimal strategy, i.e. choice of unitaries $\{V_m\}$, for a given conditional distribution $p(m|\theta)$. Here, we consider another strategy which, in general, is sub-optimal, but it works for systems with arbitrary dimension.

In this strategy we choose unitaries [3]

$$V_m = R_m = \frac{1}{\sqrt{K_m^\dagger K_m}} K_m^\dagger \ , \tag{65}$$

where

$$K_m \equiv \int d\theta \ p(\theta, m) \ U_\theta, \tag{66}$$

and $p(\theta, m) = p(m|\theta)p(\theta)$ (To simplify the discussion we have assumed operators K_m are all full-rank). Then, we prove that using these unitaries we can achieve the average entanglement fidelity $\overline{F}(\{R_m\})$, which satisfies

$$\Big[\frac{1}{d}\sum_{m\in\mathcal{M}} \| \int d\theta \ p(\theta, m) U_\theta \|_1\Big]^2 \leq \overline{F}(\{R_m\}) \leq \overline{F}_{\text{opt}} \ . \tag{67}$$

This lower bound should be compared with the upper bound $\overline{F}_{\text{opt}} \leq \frac{1}{d^2}\sum_{m\in\mathcal{M}} \Big\| \int d\theta \ p(\theta, m)[U_\theta \otimes U_\theta^*]\Big\|_1$.

It turns out that the unitaries $\{R_m\}$ defined above, are in fact the set of unitaries which maximizes the average performance quantified by

$$G_{\text{avg}}(\{V_m\}) = \int d\theta \ p(\theta) \sum_{m\in\mathcal{M}} p(m|\theta) \times \frac{1}{2d}\Big[\text{Tr}(V_m U_\theta) + \text{Tr}(V_m^\dagger U_\theta^\dagger)\Big], \tag{68}$$

Note that this formula is identical to the formula for the average entanglement fidelity, in Eq.14, except we have replaced entanglement fidelity $\frac{1}{d^2}|\text{Tr}(V_m U_\theta)|^2$ with $\frac{1}{2}\Big[\text{Tr}(V_m U_\theta) + \text{Tr}(V_m^\dagger U_\theta^\dagger)\Big]$. Let

$$G(W_1 W_2^\dagger) = \frac{1}{2d}[\text{Tr}(W_1 W_2^\dagger) + \text{Tr}(W_1^\dagger W_2)] \ , \tag{69}$$

where d is the dimension of the Hilbert space. Then,

$$G_{\text{avg}}(\{V_m\}) \equiv \int d\theta \ p(\theta) \sum_{m\in\mathcal{M}} p(m|\theta) \times G(V_m U_\theta) \tag{70}$$

$$= \sum_m \frac{1}{2d}[\text{Tr}(V_m K_m) + \text{Tr}(V_m^\dagger K_m^\dagger)]. \tag{71}$$

Putting $V_m = R_m = \frac{1}{\sqrt{K_m^\dagger K_m}} K_m^\dagger$, we find that

$$G_{\text{avg}}(\{R_m\}) = \frac{1}{d}\sum_m \text{Tr}(\sqrt{K_m^\dagger K_m}) = \frac{1}{d}\sum_m \|K_m\|_1 \ . \tag{72}$$

It can be easily seen that this is the maximum value of $G_{\text{avg}}(\{V_m\})$, for any choice of $\{V_m\}$. To see this note that for any unitary V_m,

$$|\text{Tr}(V_m K_m)| \leq \|K_m\|_1 \ , \tag{73}$$

[3] It could be interesting to consider the following definition $\tilde{K}_m \equiv \int d\theta \ \sqrt{p(\theta|m)} \ U_\theta$, and $\frac{1}{\sqrt{K_m^\dagger K_m}} K_m^\dagger$. Also, we can exploit the phase freedom in defining the unitaries U_θ.

which implies

$$G_{\text{avg}}(\{V_m\}) = \frac{1}{2d}\sum_m [\text{Tr}(V_m K_m) + \text{Tr}(V_m^\dagger K_m^\dagger)] \leq \frac{1}{d}\sum_m |\text{Tr}(V_m K_m)| \leq \frac{1}{d}\sum_m \|K_m\|_1 \ . \tag{74}$$

Finally, to prove the bound in Eq.67, we note that

$$G(W_1 W_2^\dagger) \leq \frac{1}{d}|\text{Tr}(W_1 W_2^\dagger)| = \sqrt{F_e(W_1 W_2^\dagger)}, \tag{75}$$

Then, using the fact that the square root is a concave function we find

$$G_{\text{avg}}(\{V_m\}) = \int d\theta \ p(\theta) \sum_{m\in\mathcal{M}} p(m|\theta) \times G(V_m U_\theta) \leq \int d\theta \ p(\theta) \sum_{m\in\mathcal{M}} p(m|\theta) \times \sqrt{F_e(V_m U_\theta)} \tag{76}$$

$$\leq \sqrt{\int d\theta \ p(\theta) \sum_{m\in\mathcal{M}} p(m|\theta) \times F_e(V_m U_\theta)} \ . \tag{77}$$

We conclude that for any set of unitaries $\{V_m\}$, it holds that

$$G^2_{\text{avg}}(\{V_m\}) \leq \overline{F}(\{V_m\}) \ . \tag{78}$$

Therefore, for $V_m = R_m$, this implies

$$\frac{1}{d}\sum_m \|K_m\|_1 = G^2_{\text{avg}}(\{R_m\}) \leq \overline{F}(\{R_m\}) \ \leq \overline{F}_{\text{opt}} \ . \tag{79}$$

[1] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge Series on Information and the Natural Sciences (Cambridge University Press, 2000).

[2] M. Horodecki, P. Horodecki, and R. Horodecki, Physical Review A **60**, 1888 (1999).

[3] M. A. Nielsen, Physics Letters A **303**, 249 (2002).