Estimation of \( n \) non-identical unitary channels

Caleb O’Loan

School of Mathematics and Statistics, University of St Andrews, KY16 9SS

(Dated: February 11, 2008)

We investigate the simultaneous estimation of \( n \) not necessarily identical unitary channels using multi-partite entanglement. We examine whether it is possible for the rate at which the mean square error decreases to be greater than that using the channels individually. For a reasonably general situation, in which there is no functional dependence between the channels, we show that this is not possible. We look at a case in which the channels are not necessarily identical but depend on a common variable. In this case, the mean square error decreases more rapidly using multi-partite entanglement.

PACS numbers: 03.65.Ta, 03.67.-a

Introduction. Estimation of quantum states and channels is of fundamental importance to quantum information theory. Estimation of unitary channels, when \( n \) copies are available, has received a lot of attention; for an accessible overview see [1]. Many schemes have been devised for which the error decreases at a much faster rate, compared to a straightforward approach of using each channel separately [2, 3, 4, 5, 6]. This is analogous to the problem of transmitting a reference frame [1, 2, 3, 4, 5]. This increase in the rate of estimation is possible both with [2, 3, 4, 5, 6] and without [7, 8] the use of entanglement.

As far as we are aware, no work has been done on the situation in which there are \( n \) non-identical channels. In this paper we investigate whether an increase in the rate of estimation is possible in this case. We look at the following two cases: (i) there are \( n \) unitary channels belonging to \( SU(d) \) specified by the parameters \( \theta^1, \ldots, \theta^n \), (ii) there are \( n \) not necessarily identical one-parameter unitary channels which depend on a common variable. In the former scheme an increase in the rate of estimation is not possible; in the latter scheme it is.

To quantify the performances of estimation schemes we use the asymptotic limit of the mean square error \( E[(\hat{\theta}^{(t)} - \theta^{(t)})^2] \). Using the maximum likelihood estimator, as the number of measurements \( N \to \infty \), the mean square error is approximately \( (1/N)F^{-1}_M(\theta) \) [3], where \( F_M(\theta) \) is the Fisher information matrix obtained from a single measurement \( M \). The Fisher information for the parameter \( \theta = (\theta^1, \ldots, \theta^n) \) is the \( p \times p \) matrix with entries

\[
F_{ij}(\theta) = \int p(\xi) \frac{\partial}{\partial \theta^j} \ln p(\xi; \theta) \frac{\partial}{\partial \theta^k} \ln p(\xi; \theta) d\xi.
\]

The Cramér–Rao inequality states that the mean square error of an unbiased estimator \( \hat{\theta}(x) \) is less than or equal to the inverse of the Fisher information,

\[
\text{m.s.e.}[\hat{\theta}(x)] \geq F_M(\theta)^{-1}.
\]

It has been shown [10] that the SLD quantum information \( H(\theta) \) is an upper bound on the Fisher information, i.e. \( F_M(\theta) \leq H(\theta) \). The SLD quantum information for the parameter \( \theta = (\theta^1, \ldots, \theta^n) \) is defined as the matrix with entries

\[
H_{jk} = \Re \text{tr} \{ \lambda^j \rho \lambda^k \},
\]

where \( \lambda^j \) is any self-adjoint solution to the matrix equation

\[
\frac{d\rho}{dt} = \frac{1}{2} (\rho \lambda^j + \lambda^j \rho).
\]

Given a state \( \rho_\theta \) depending on some unknown parameter \( \theta \), and a POVM \( \{M_m\} \) we get a measurement outcome \( x \). The quantum Cramér–Rao inequality states that the mean square error of an unbiased estimator \( \hat{\theta}(x) \) of \( \theta \) is less than or equal to the inverse of the SLD quantum information, i.e.

\[
\text{m.s.e.}[\hat{\theta}(x)] \geq H^{-1}(\theta).
\]

When \( p \geq 2 \), there exist families of states \( \rho_\theta \) and \( \sigma_\theta \) for which neither \( H(\rho_\theta) \geq H(\sigma_\theta) \) or \( H(\rho_\theta) \leq H(\sigma_\theta) \) is true. To deal with this, we compare the traces \( \text{tr}\{H(\theta)\} \) of the SLD quantum information.

Previous work has looked at how the error scales with the number of times \( U \) is used. We look at how the error scales with the number of input states used, as this is more convenient for us.

I. ‘INDEPENDENT’ CHANNELS

First we look at estimating \( n \) unitary channels from \( SU(d) \). The \( j \)th channel is specified by the parameter \( \theta^j = (\theta^j_1, \ldots, \theta^j_{d-1}) \). These channels are supposed ‘independent’ in that there is no functional relationship between \( \theta^1, \theta^2, \ldots, \theta^{n-1} \) and \( \theta^n \). Estimation of the channels is a parametric problem.

We look at the mapping

\[
\rho_\theta \mapsto \otimes_{j=1}^n (U_{\theta^j} \otimes \mathbb{I}_R) \rho_\theta (U_{\theta^j}^\dagger \otimes \mathbb{I}_R).
\]
We show in Appendix A that using a tensor product of maximally entangled input states $\rho_0 = \otimes_{j=1}^{n} \rho_{mes}^{j}$ is sufficient to maximize the trace of the SLD quantum information of the output states of (4). This is not a necessary condition, as the $2n$-partite entangled state $1/\sqrt{2} \sum_{i} |e_i \rangle \otimes \cdots \otimes |e_i \rangle$ also attains the maximum SLD quantum information. However, since these states are significantly harder to produce, we are better off using maximally entangled states. The SLD quantum information using a tensor product of maximally entangled states is attainable (1), so asymptotically the mean square error is $(1/N)H^{-1}(\theta)$. 

For the mapping (1), an optimal estimation procedure for $n$ ‘independent’ unitary channels, in terms of $\text{tr}\{H\}$, is to estimate each one individually using a maximally entangled input state. The optimality of this procedure, in terms of $\text{tr}\{H(\theta)\}$ has been shown by Ballester (1).

II. ‘DEPENDENT’ CHANNELS

We look at the case where we have $n$ not necessarily identical channels which depend on a common parameter $\theta$. These channels are of the form

$$U_{\theta}^1 = \begin{pmatrix} 1 & 0 \\ 0 & e^{i f_1(\theta)} \end{pmatrix}, \ldots, \ U_{\theta}^n = \begin{pmatrix} 1 & 0 \\ 0 & e^{i f_n(\theta)} \end{pmatrix},$$

(5)

where $0 \leq \theta \leq t$ and $\mathcal{H} = \mathbb{C}^2$. We can use each channel $N$ times. We impose the following conditions on the functions $f_j$: (a) $f_j(\theta) : \mathbb{R} \to \mathbb{R}$, (b) $df_j(\theta)/d\theta \geq 0$, and (c) $0 \leq \sum_j f_j(\theta) \leq \pi$, for all $j$ and $\theta$. We look at the mapping

$$\rho_0 \mapsto \otimes_{j=1}^{n} (U_{\theta}^j \otimes \mathbb{I}_R) \rho_0 (U_{\theta}^j \dagger \otimes \mathbb{I}_R),$$

(6)

where $\rho_0 \in \mathcal{S}(\mathbb{C}^2 \otimes \mathcal{H}_R)$; we denote by $\mathcal{S}(\mathcal{H})$ the set of states on $\mathcal{H}$. We compare the SLD quantum information (ii) using a tensor product of maximally entangled states, (ii) using a $2n$-partite entangled state, in this case using $2n$-partite entanglement gives considerably larger SLD quantum information.

Using the input state $\rho_0 = \rho_{mes}^1 \otimes \cdots \otimes \rho_{mes}^n$, where $\rho_{mes}^j = |\psi_u^j \rangle \langle \psi_u^j|$, $|\psi_u^j \rangle = 1/\sqrt{2}(|00 \rangle + |11 \rangle)$, gives an SLD quantum information of $\sum_{j=1}^{n} (df_j/d\theta)^2$, which is attainable using a tensor product of the POVM $\{M_0^j = |\psi_1^j \rangle \langle \psi_1^j|, M_1^j = \mathbb{I} - |\psi_u^j \rangle \langle \psi_u^j|\}$. We use the input state $\rho_0 = |\psi_0 \rangle \langle \psi_0|$, where $|\psi_0 \rangle = 1/\sqrt{2}(|00 \rangle + |11 \rangle)$. We get an SLD quantum information of $\sum_{j=1}^{n} (df_j/d\theta)^2$. As $df_j/d\theta \geq 0$ for all $j$, this is considerably larger than the SLD quantum information using a tensor product of maximally entangled states.

The phase $\phi$ of the output state is in one-to-one correspondence with $\theta$, because of conditions (b) and (c); in fact $\phi = \sum_{j=1}^{n} f_j(\theta)$. Using the POVM $\{M_0 = |\psi_0 \rangle \langle \psi_0|, M_1 = \mathbb{I} - |\psi_0 \rangle \langle \psi_0|\}$ we get $p(0; \phi) = \cos^2(\phi/2)$. We perform $N$ times and each time use this POVM. From the measurement outcomes we get an estimate $\hat{\phi}$ of $\phi$ by $\cos^2(\phi/2) = n_0/N$, where $n_0$ is the number of times we get the outcome $x = 0$. Because of condition (c) on the functions $f_j(\theta)$, $\cos^2(\phi/2)$ is in one-to-one correspondence with $\hat{\phi}$ and hence $\hat{\theta}$. This POVM gives a Fisher information equal to the SLD quantum information. Hence as $N \to \infty$ we get a mean square error of $1/(N(\sum_{j=1}^{n} (df_j/d\theta)^2))$ to $1/(N(\sum_{j=1}^{n} (df_j/d\theta)^2))$ using a tensor product of maximally entangled states.

It is known that when we have $n$ identical simple unitary channels, we can obtain an increase in the rate of estimation without using entanglement (1). A simple way is to use each of the $n$ channels in sequence on a single input state. Consider the unitary channel $U_\theta = \text{Diag}(1, e^{i\theta})$, i.e.

$$\rho_0 \mapsto \rho_0^0 U_\theta^n, \ U_\theta^n = U_\theta U_\theta \cdots U_\theta.$$

(7)

We repeat $N$ times using the input state $\rho_0 = |\psi_x \rangle \langle \psi_x|$, where $|\psi_x \rangle = 1/\sqrt{2}(|0 \rangle + |1 \rangle)$, and the POVM $\{\{0\} = |\psi_x \rangle \langle \psi_x|, \{1\} = |\psi_x \rangle \langle \psi_x|\}$. We call this the sequential scheme. As $N \to \infty$, we obtain a mean square error that scales as $1/(Nn^2)$. We can use this sequential scheme with $n$ non-identical channels, i.e.

$$\rho_0 \mapsto (U_\theta^3 \cdots U_\theta^0) \rho_0 (U_\theta^0 \cdots U_\theta^3)^\dagger,$$

(8)

where $\rho_0 \in \mathcal{S}(\mathcal{H})$. Using the input state $\rho_0 = |\psi_x \rangle \langle \psi_x|$, where $|\psi_x \rangle = 1/\sqrt{2}(|0 \rangle + |1 \rangle)$, and the POVM $\{\{0\} = |\psi_x \rangle \langle \psi_x|, \{1\} = |\psi_x \rangle \langle \psi_x|\}$ we get $p(0; \phi) = \cos^2(\phi/2)$. Performing $N$ times we get an estimate $\hat{\phi}$ and hence $\hat{\theta}$. We get the same Fisher information as (6) using a 2n-partite entangled state, and hence the same mean square error.

Since multi-partite entanglement is difficult to create, we are better off using the sequential scheme (8).

Conditions (b) and (c) for the functions $f_j(\theta)$ are very strict. We can still get a considerable increase in the rate of estimation without these conditions. If condition (c) does not hold, the phase $\phi$ of the output state is not in one-to-one correspondence with $p(0; \phi) = \cos^2(\phi/2)$. If we modify condition (c) to $0 \leq \sum_j f_j(\theta) \leq 2\pi$, we can still find $\phi$ but we shall need to perform an extra measurement $\{\psi_y \rangle \langle \psi_y|, \mathbb{I} - |\psi_y \rangle \langle \psi_y|\}$, where $|\psi_y \rangle = 1/\sqrt{2}(|0 \rangle + i|1 \rangle)$, a small number of times to determine the sign of $\cos(\phi/2)$. Then we can estimate $\phi$ and hence $\theta$. If $0 \leq \sum_j f_j(\theta) \leq 2\pi$ does not hold, the phase $\phi$ of the output state is not in one-to-one correspondence with $\theta$.

We can get around this using a method similar to that of Zhengfeng et al (4), based on Rudolph and Grover (3).

Zhengfeng et al (1) looked at the case where $f_j(\theta) = 2\pi \theta$ for all $j$. Their scheme involves first using a single channel $n$ times to get an interval in which $\theta$ almost certainly lives. Then they use two or three channels in sequence to ‘amplify’ $\theta$. This is repeated $n$ times until they get a narrower interval for $\theta$. This is continued until all channels are being used simultaneously. The mean square error scales as $(\log N'/N'')^2$, where $N'$ is the total number of times $U_\theta$ is used.
In our case, the situation is more complex, as the functions $f_j(\theta)$ are more general and not necessarily identical. The finer details of how we go about this and how the mean square error would behave, depend on the functions $f_j(\theta)$. We give a very brief overview of a possible procedure. We start by using a single unitary $U_1$ $n$ times to get an interval for $f_1(\theta)$ and hence $\theta$. Then we use $U_2U_1$ to get an estimate of $f_2(\theta)$ and hence a more accurate estimate of $\theta$. We continue this process till we are using all the channels simultaneously and we have a very narrow interval for $\theta$. We expect that asymptotically, the mean square error is approximately $1/(N(\sum_j df_j/d\theta)^2)$, where $N$ is the number of input states used. We leave a more in-depth analysis for further work.

If neither (b) nor (c) are satisfied we propose the following scheme: (i) Use each of the channels individually $n$ times to get an estimate $\hat{\theta}$, (ii) Divide the channels into two groups: $A = \{U_\theta^j, df_j/d\theta \geq 0 \text{ at } \hat{\theta}\}, B = \{U_\theta^j, df_j/d\theta < 0 \text{ at } \hat{\theta}\}$, (iii) Use an iterative procedure for the two groups separately, but sharing information about $\theta$ to make the confidence intervals shorter.

**Appendix A: Proof**

We are looking at unitary channels of the form $U_\theta = \exp(i \sum_j \theta_j t_j)$ where $t_j = t_j^+$, $\text{tr}(t_j) = 0$ and $\text{tr}(t_j^t) = \delta_{jk}$. We denote by $H_{\theta}(\rho_0)$, the SLD quantum information for the $j$th channel

$$
\rho_0 \mapsto (U_\theta \otimes 1)\rho_0(U_\theta^\dagger \otimes 1).
$$

The SLD quantum information of (4) using a tensor product of maximally mixed states is $H_{\theta}(\otimes_j^m \rho_{\text{mes}}) = \text{Diag}(H_{\theta_1}(\rho_{\text{mes}}), \ldots, H_{\theta_n}(\rho_{\text{mes}}))$.

**Lemma 1** For all unitary channels of the form

$$
\rho_0 \mapsto (U_\theta \otimes 1)\rho_0(U_\theta^\dagger \otimes 1), \tag{A1}
$$

the trace of the SLD quantum information is maximized by a maximally mixed state, i.e.

$$
\text{tr}\{H_{\theta}(\rho_0)\} \leq \text{tr}\{H_{\theta}(\rho_{\text{mes}})\}. \tag{A2}
$$

Equality holds in (A2) if and only if $\rho_0$ is a maximally mixed state.

**Proof.** From the appendix of Ballester [12] we know that for the channel (A1)

$$
\text{tr}\{(H_{\theta}(\rho_{\text{mes}}))^{-1}H_{\theta}(\rho_0)\} \leq d^2 - 1. \tag{A3}
$$

It is simple to show that for unitary channels of the form $U_\theta = \exp(i \sum_j \theta_j t_j)$ we have $H_{\theta}(\rho_{\text{mes}}) = (4/d)I_{d^2-1}$. Substituting into (A3) we get

$$
\text{tr}\{H_{\theta}(\rho_0)\} \leq \frac{4(d^2 - 1)}{d} = \text{tr}\{H_{\theta}(\rho_{\text{mes}})\}. \tag{A4}
$$

Since equality holds in (A3) if and only if $\rho_0$ is a maximally entangled state [12], equality holds in (A4) if and only if $\rho_0$ is a maximally mixed state.

**Lemma 2** The trace of the SLD quantum information for (4) is maximized by a tensor product of maximally mixed states, i.e.

$$
\text{tr}\{H_{\theta}(\rho_0)\} \leq \text{tr}\{H_{\theta}(\otimes_j^n \rho_{\text{mes}})\}, \tag{A5}
$$

$\rho_0 \in S(\otimes_i^m (\mathcal{H} \otimes \mathcal{R}))$, $\rho_{\text{mes}} \in S(\mathcal{H} \otimes \mathcal{R})$.

For pure states a solution of (2) is $\lambda^2 = 2d\rho_0/d\theta$. It is not difficult to show that for the set of states $U_\theta \rho_0 U_\theta^\dagger$ the SLD quantum information is the matrix with entries

$$
H_{\theta_{jk}}(\rho_0) = 4\text{Re}(U_{\theta_{jk}}^\dagger \rho_0 U_{\theta_{jk}}^\dagger) + 4\left(\text{tr}(U_{\theta_{jk}}^\dagger \rho_0 U_{\theta_{jk}})\right)^2,
$$

where $U_{\theta_{jk}} = \partial_{\theta_{jk}} U_\theta$. In (A3)

$$
H_{\theta_{mm}}(\rho_0) = 4\text{tr}(U_{\theta_{jk}}^\dagger \rho_0 U_{\theta_{jk}}^\dagger) + 4\left(\text{tr}(U_{\theta_{jk}}^\dagger \rho_0 U_{\theta_{jk}})\right)^2,
$$

Consider an arbitrary diagonal element of the SLD quantum information, i.e. $H_{\theta_{mm}}$, where $m(j,k) = (j - 1)p + k$, corresponding to the parameter $\theta_{jk}$. From (A6),

$$
H_{\theta_{mm}}(\rho_0) = 4\text{tr}(U_{\theta_{jk}}^\dagger \rho_0 U_{\theta_{jk}}^\dagger) + 4\left(\text{tr}(U_{\theta_{jk}}^\dagger \rho_0 U_{\theta_{jk}})\right)^2,
$$

by (A5),

$$
H_{\theta_{mm}}(\rho_0) = \sum_{i=1}^d p_i H_{\theta_i}(\rho_i)_{kk}.
$$

Summing over $k$ we get

$$
\sum_k H_{\theta_{m(j,k)=m(j,k)}}(\rho_0) = \sum_{k=1}^d \sum_{i=1}^d p_i H_{\theta_i}(\rho_i)_{kk} \leq \text{tr}\{H_{\theta}(\rho_{\text{mes}})\}.
$$
Summing over $j$ we get \([A_{ij}]\).

[1] J. Zhengfeng, G. Wang, R. Duan, Y. Feng, and M. Ying, quant-ph/0610060 (2006).
[2] M. Hayashi, Phys. Rev. Lett. 354, 183 (2006).
[3] J. Kahn, Phys. Rev. A 75, 022326 (2007).
[4] E. Bagan, M. Baig, and R. Munoz-Tapia, Phys. Rev. A 69, 050303 (2004).
[5] E. Bagan, M. Baig, and R. Munoz-Tapia, Phys. Rev. A 70, 030301 (2004).
[6] G. Chiribella, G. M. D’Ariano, P. Perinotti, and M. Sacchi, Phys. Rev. Lett. 93, 180503 (2004).
[7] T. Rudolph and L. Grover, Phys. Rev. Lett 91, 217905 (2003).
[8] M. de Burgh and S. Bartlett, Phys. Rev. A 72, 042301 (2005).
[9] M. Hayashi, Quantum Information: an introduction (Springer, 2006).
[10] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. 72, 3439 (1994).
[11] M. Ballester, Phys. Rev. A 69, 022303 (2004).
[12] M. Ballester, Phys. Rev. A 70, 032310 (2004).