EPJ manuscript No.
(will be inserted by the editor)

Unlearning Quantum Information

Daniel K. L. Oi

SUPA Department of Physics, University of Strathclyde, Glasgow G4 0NG, United Kingdom

the date of receipt and acceptance should be inserted later

Abstract. Measurement can drive quantum dynamics, for example in ancilla driven quantum computation where unitary evolution is generated by measurements that extract no information. Where a measurement does reveal some information about the system, it may sometimes be possible to "unlearn" this information and restore unitary evolution through subsequent measurements. Here we analyse two methods of quantum "unlearning" and present a simplified proof of the bound on the probability of successfully applying the required correction operators. The probability of successful recovery is inversely related to the ability of the initial measurement to exclude the possibility of a state. As a consequence there exist unrecoverable measurements that provide little information gain.

PACS. 03.65.Aa Quantum systems with finite Hilbert space – 03.65.Ta Foundations of quantum mechanics; measurement theory – 03.65.-a Quantum information – 03.67.Pp Quantum error correction and other methods for protection against decoherence

1 Introduction

Quantum information processing schemes such as measurement based quantum computation [1], ancilla driven quantum computation [2–4], and holonomic degenerate projections [5,6], drive unitary quantum dynamics by measurements that learn nothing about the system. A non-ideal operation may gain information but this can sometimes be "unlearned" by subsequent conditional measurements to restore unitary evolution. Previously this has been studied in the context of the theory of reversing measurement [7–10], together with experimental proposals [11–13] and demonstrations [14,15].

Here, we present a simplified proof of the bound on the success probability of such corrective measures. In addition to conventional Procrustean filtering, we analyse the asymptotic success probability of partial filtering for measurement reversal. As the correction probability is related to the minimum eigenvalues of the measurement operators, this leads to a class of uncorrectable measurements that nonetheless provide little information gain.

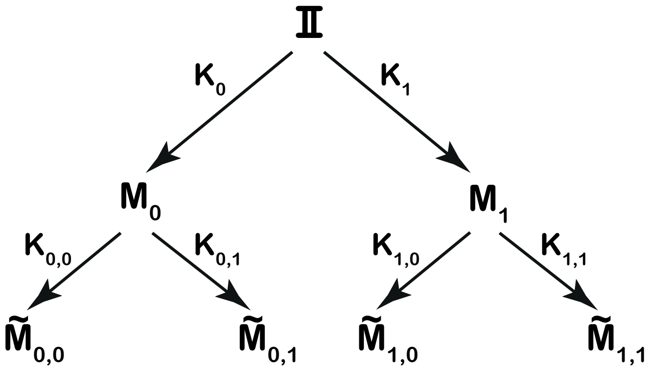

Fig. 1. Two-level binary POVM tree. Each bifurcation represents a binary POVM with two Kraus operators labeling the arrows. The nodes of the tree represent the cumulative measurement operator corresponding to the sequence of results leading to that node. The cumulative Kraus operator $\widetilde{K}_{j,k}$ consists of the product of the Kraus operators along the path down the branch. The sum of the $\widetilde{M}_{j,k}$ children of a branch sum up to parent node, $M_j = \widetilde{M}_{j,0} + \widetilde{M}_{j,1}$.

2 Preliminaries

A generalized measurement, or positive operator valued measure (POVM), can be described by a set of positive operators $\{M_j\}$ summing to the identity, $\sum_j M_j = \mathbb{I}$. The probability of obtaining outcome j when measuring a system described by a density operator ρ is $p_j = \mathrm{Tr}[M_j\rho]$. The post-measurement state is not uniquely defined by M_j in general, but is given by $\rho_j = \frac{K_j\rho K_j^\dagger}{\mathrm{Tr}[K_j^\dagger K_j\rho]}$ where $M_j = K_j^\dagger K_j$, and $\{K_j\}$ are Kraus operators. We will consider operations where the post-measurement state has the same dimension as the input.

A cascaded sequence of measurements (Fig. 1) results in a cumulative Kraus operator that is the product of the individual Kraus operators associated with each sequential result [16]. I.e. An initial measurement has Kraus operators $\{K_j\}$ and conditional on the result j a second measurement is performed with Kraus operators $\{K_{j,k}\}$, the cumulative Kraus operator associated with joint result j

and then k is given by $\widetilde{K}_{j,k} = K_{j,k}K_j$, and the POVM element is given by $\widetilde{M}_{j,k} = (\widetilde{K}_{j,k})^\dagger \widetilde{K}_{j,k}$.

We will consider the case where ideally we would like the Kraus operators to be proportional to a unitary, $K_j = q_j U_j$ where $0 < q_j \leq 1$ for some unitary U_j. This results in $M_j = q_j^2 \mathbb{I}$, hence the measurement probabilities q_j^2 are independent of ρ, i.e. obtaining outcome j reveals no information about the state of the system. This ensures that the conditional evolution is unitary, $\rho_j = U_j \rho U_j^\dagger$.

In ancilla-driven quantum computation (ADQC) [2–4], the coupling and measurement of an ancilla qubit to the system results in a two-outcome POVM with unitary Kraus operators that are related by a Pauli correction. This requires the coupling between system and ancilla to be of a special form, and that the ancilla qubit be prepared and measured in particular directions [3]. Otherwise the effective Kraus operators may not result in the desired unitary conditional evolution but may reveal information about the system.

We will find useful the singular value decomposition (SVD) for the Kraus operators, $K_j = V_j D_j W_j$, where V_j and W_j are unitaries and $D_j = \mathrm{diag}(q_j^r)$ with non-negative singular values q_j^r. A necessary and sufficient condition for K_j to represent a conditional unitary is that the singular values be uniform i.e. $q_j^r = q_j \ \forall r$ and $D_j = q_j \mathbb{I}$. Deviations from uniformity in the singular values represents information gain as the probability for obtaining the outcome j would be state dependent. Since the singular values encodes the nonunitary behaviour and to simplify the analysis in the rest of the paper, we will assume that all Kraus operators have aligned pre- and post-measurement bases [17], this is equivalent to setting $V_j = W_j = \mathbb{I}$ for all SVDs and the Kraus operators being diagonal. We can consider the general case where corrective Kraus operators do not have aligned bases to the initial measurement but we see later that this does not alter the basic results.

3 Procrustean Filtering

We can correct an initial non-unitary inducing measurement by filtering similar to that used for entanglement concentration [18] [1]. Assume for the first measurement we obtain the outcome associated with Kraus operator $K_0 = \mathrm{diag}(q_0^r)$ where the q_0^r are not all the same. The probability of this result is $p_0 = \mathrm{Tr}[K_0^\dagger K_0 \rho]$ and varies from $(q_0^{r_{min}})^2 \leq p_0 \leq (q_0^{r_{max}})^2$ depending on the state, hence we gain information and the evolution is non-unitary. We now try to correct the evolution with another measurement with diagonal Kraus operators $K_{0,k} = \mathrm{diag}(q_{0,k}^r)$, $k = 0, 1$ resulting in the conditional cumulative Kraus operators, $\widetilde{K}_{0,k} = K_{0,k}K_0 = \mathrm{diag}(q_{0,k}^r q_0^r)$. We may set the singular values so that for one of the outcomes we restore unitary evolution.

[1] The Procustrean Method of entanglement distillation refers to the Greek legend of Procrustes, a sadistic host who would cut off the legs of guests who were too tall for their beds. Procrustean filtering removes parts of the wavefunction to better match the desired state.

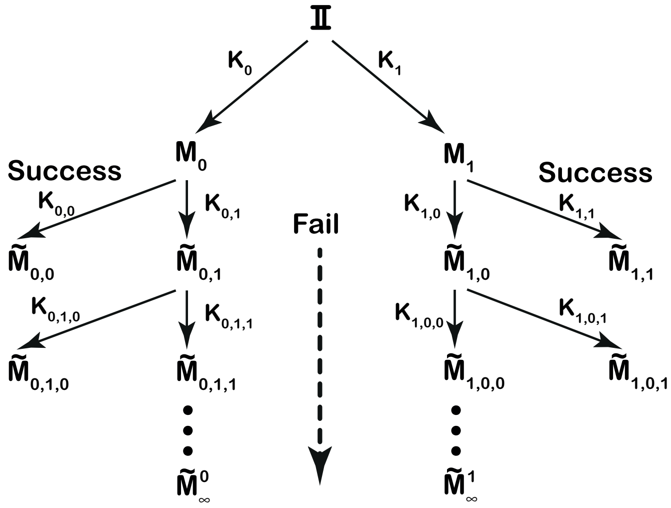

Fig. 2. Partial Filtering. Instead of succeeding or failing outright after one step, we can partially filter out corrected portions of the evolution, represented by the paths leading out to the sides. The vertical downward arrows represent partial failures, upon which we can retry recovery. The failure probability is given by the sum of the limiting residual cumulative measurement operators, $p_{fail}\mathbb{I} = \widetilde{M}_\infty^0 + \widetilde{M}_\infty^1$.

Let us choose the Kraus operator $K_{0,0}$ to correct K_0. If $q_0^{r_{min}}$ is the smallest singular value of K_0, then setting $q_{0,0}^r = q_0^{r_{min}}/q_0^r$ results in the cumulative operation $\widetilde{K}_{0,0} = K_{0,0}K_0 = q_0^{r_{min}}\mathbb{I}$ where the probability of this branch is given by $(q_0^{r_{min}})^2$ independent of the initial state as required by unitarity. We note that the maximum singular value of $K_{0,0}$ is 1 which means that the other outcome $K_{0,1}$ will have at least one vanishing singular value due to completeness of the POVM $\{M_{0,k}\}$. Hence $\widetilde{K}_{0,1}$ will have a non-trivial nullspace and it will be impossible to further correct this branch of the measurement tree.

If at the first measurement we obtained the complementary result $K_1 = \mathrm{diag}(q_1^r)$, then a subsequent correction would result in outcome $\widetilde{K}_{1,1} = q_1^{r_{min}}\mathbb{I}$ with probability $p_{1,1} = (q_1^{r_{min}})^2$. The completeness of the measurement $\{M_j\}$ implies that $(q_1^{r_{min}})^2 = 1 - (q_0^{r_{max}})^2$, hence the total probability of a successful correction after the initial measurement is $p_{tot} = 1 - [(q_0^{r_{max}})^2 - (q_0^{r_{min}})^2]$, or one minus the visibility. We can generalize the result to the case where the first measurement has more than two outcomes but it is still sufficient for each correction to be a binary POVM. In this case, the maximum success probability is given by $p_{tot} = \sum_j (q_j^{r_{min}})^2$. Hence the uncorrectable non-unitary action of the initial measurement is determined by how much an each outcome *excludes* a state compared with others.

4 Partial Filtering

We saw in the above section that we can choose our corrective measurements to either succeed, or fail entirely with no further recourse. An alternate strategy would be to succeed on one outcome, but the alternative could still be further correctable. We shall illustrate this in the case of a single qubit system.

Let the initial binary outcome measurement have Kraus operators, $K_0 = \mathrm{diag}(a, b)$ and $K_1 = \sqrt{\mathbb{I} - K_0^\dagger K_0}$ where $1 > a > b > 0$. Supposing we obtain outcome K_0, we could correct the evolution using the method in the previous section or alternatively we can choose, for example, the partial filtering operators $K_{0,0} = \mathrm{diag}(b, a)$ and $K_{0,1} = \sqrt{\mathbb{I} - K_{0,0}^\dagger K_{0,0}}$. In the case of the result $K_{0,0}$, we achieve the cumulative evolution $\widetilde{K}_{0,0} = ab\mathbb{I}$, but the unsuccessful outcome $\widetilde{K}_{0,1}$ still has full rank and could be further processed. The situation then reduces to that of before but with a new effective Kraus operator $\widetilde{K}_{0,1} = \mathrm{diag}(a^{(1)} = a\sqrt{1-b^2}, b^{(1)} = b\sqrt{1-a^2})$ and we can try to apply another round of corrections as shown in Fig. 2.

This gives a recursive formula for the success probability for the K_0 branch,

$$p_{tot}^2 = \sum_j p_j^0, \quad p_j^0 = \left(a^{(j)} b^{(j)}\right)^2,$$
$$a^{(j+1)} = a^{(j)}\sqrt{1 - b^{(j)2}}, \quad b^{(j+1)} = b^{(j)}\sqrt{1 - a^{(j)2}}, \quad (1)$$

and for the K_1 branch,

$$p_{tot}^1 = \sum_j p_j^1, \quad p_j^1 = \left(c^{(j)} d^{(j)}\right)^2,$$
$$c^{(j+1)} = c^{(j)}\sqrt{1 - d^{(j)2}}, \quad d^{(j+1)} = d^{(j)}\sqrt{1 - c^{(j)2}}, \quad (2)$$

where $c = \sqrt{1-a^2}$ and $d = \sqrt{1-b^2}$, and $a^{(0)} = a$ *etc.* It is simple to check that $a^{(j)} = d^{(j)}$ and $b^{(j)} = c^{(j)}$ $\forall j \geq 1$.

In order to compute the limiting value of the success probability, it is easier to compute the probability of failure. This can be found by finding the $j \to \infty$ limit of the unsuccessful Kraus operators given by

$$\widetilde{K}_\infty^0 = \mathrm{diag}(a^{(\infty)}, b^{(\infty)}),$$
$$\widetilde{K}_\infty^1 = \mathrm{diag}(c^{(\infty)}, d^{(\infty)}), \quad (3)$$

and the total failure probability is

$$p_{fail}\mathbb{I} = \widetilde{M}_\infty^0 + \widetilde{M}_\infty^1 = \left(a^{(\infty)2} + b^{(\infty)2}\right)\mathbb{I}, \quad (4)$$

where $\widetilde{M}_\infty^{0,1} = (\widetilde{K}_\infty^{0,1})^\dagger \widetilde{K}_\infty^{0,1}$

To solve the recursion formula, we first note that

$$a^{(j+1)2} - b^{(j+1)2} = a^{(j)2} - b^{(j)2} = a^2 - b^2. \quad (5)$$

We also note that the fixed points of the recursion relation are when $b^{(\infty)} = 0$ leading to the limit

$$\widetilde{K}_\infty^0 = \mathrm{diag}(\sqrt{a^2 - b^2}, 0),$$
$$\widetilde{K}_\infty^1 = \mathrm{diag}(0, \sqrt{a^2 - b^2}), \quad (6)$$

hence $p_{fail} = a^2 - b^2$ and

$$p_{tot} = 1 - (a^2 - b^2), \quad (7)$$

which is the same as for the Procrustean method.

5 Success Bound

We present a simplified proof of the maximum probability of "unlearning" information gained by a POVM $\{M_j\}$ whose associated Kraus operators do not all produce unitary evolution [8,10]. For outcome j, we apply a corrective POVM $\{M_{j,k}\}$ where some of the outcomes $\{k'\}$ succeed, $\widetilde{M}_{j,k'} = p_{j,k'}\mathbb{I}$ and these sum to

$$\sum_{k \in \{k'\}} \widetilde{M}_{j,k} = \left(\sum_{k \in \{k'\}} p_{j,k}\right)\mathbb{I} = p_{succ}^j \mathbb{I}. \quad (8)$$

Since

$$M_j = \sum_k \widetilde{M}_{j,k} = \sum_{k \in \{k'\}} \widetilde{M}_{j,k} + \sum_{k \notin \{k'\}} \widetilde{M}_{j,k}, \quad (9)$$

the failure branches sum to

$$\sum_{k \notin \{k'\}} \widetilde{M}_{j,k} = M_j - p_{succ}^j \mathbb{I}. \quad (10)$$

Since this is a positive operator, $p_{succ}^j \leq p_j^{min} = (q_j^{r_{min}})^2$, where p_j^{min} is the minimum eigenvalue of M_j. When equality holds, the sum of the failure branch measurement operators is rank deficient and hence no longer correctable. Considering all of the branches of the initial measurement $\{M_j\}$, the maximum total probability of recovery is given by $p_{max} = \sum_j p_j^{min}$ and the Procrustean method saturates this bound by construction.

Previous work has related this bound on measurement reversal probability to the maximum information gain [10],

$$G_{max} = \frac{1}{d(d+1)}\left[d + \sum_{j=0}^{N-1} p_j^{max}\right] \quad (11)$$

for a d-dimensional system when measured by a POVM with N-outcomes whose elements each have maximum eigenvalue p_j^{max}. The quantity G_{max} is the maximum estimation fidelity averaged over all outcomes of the measurement and obeys the inequality $\frac{1}{d} \leq G_{max} \leq \frac{2}{d+1}$, where the lower and upper limits are saturated by conditional unitary evolution and complete projections respectively. In Ref. [10] the following trade-off between information gain and reversibility was demonstrated,

$$d(d+1)G_{max} + (d-1)p_{max} \leq 2d, \quad (12)$$

both unitary evolution and complete projections saturate this bound.

We can also consider the converse and ask what measurements display the greatest gap in the inequality. For unrecoverable measurements, the minimum eigenvalue of each POVM operator must be 0, for example operators of the form $M_j = q_j^2(\mathbb{I} - |\psi_j\rangle\langle\psi_j|)$. If $\{|\psi_j\rangle\}$ is a d-dimensional orthonormal basis, then $\left\{M_j = \frac{\mathbb{I} - |\psi_j\rangle\langle\psi_j|}{d-1}\right\}$ is an unrecoverable measurement but the information gain $G_{max} = \frac{d}{d^2-1}$ tends to the lower bound $\frac{1}{d}$ as $d \to \infty$.

6 Application to probabilistic teleportation

We apply the results to the well studied problem of probabilistic quantum teleportation as an illustration. Alice and Bob share a non-maximally entangled state of the form $|\Psi(\theta)\rangle = \cos\frac{\theta}{2}|00\rangle + \sin\frac{\theta}{2}|11\rangle$ where $0 < \theta < \pi/2$. Charlie gives Alice a qubit in the state $|\phi\rangle = \alpha|0\rangle + \beta|1\rangle$ to teleport to Bob with the proviso that it either arrives with unit fidelity, or else it fails. The standard solution [19] is for Alice to measure in a non-maximally entangled basis,

$$|\Psi^0\rangle = \sin\frac{\theta}{2}|00\rangle + \cos\frac{\theta}{2}|11\rangle, |\Psi^1\rangle = \sin\frac{\theta}{2}|10\rangle + \cos\frac{\theta}{2}|01\rangle$$
$$|\Psi^2\rangle = \cos\frac{\theta}{2}|00\rangle - \sin\frac{\theta}{2}|11\rangle, |\Psi^3\rangle = \cos\frac{\theta}{2}|10\rangle - \sin\frac{\theta}{2}|01\rangle.$$

The first two outcomes $\{|\Psi^0\rangle, |\Psi^1\rangle\}$ each occur with probability $\frac{1}{4}\sin^2\theta$ and result in Pauli correctable unitary quantum channels between Alice and Bob.

For the other two results, Alice obtains some information about $|\phi\rangle$ resulting in operations with singular values $\{\cos^2\frac{\theta}{2}, \sin^2\frac{\theta}{2}\}$. Bob can choose to reverse the non-unitary dynamics by filtering with probability $\sin^4\frac{\theta}{2}$ in both cases. The total probability of Alice and Bob to succeed in teleporting $|\phi\rangle$ is $p = 2\left(\frac{1}{4}\sin^2\theta\right) + 2\left(\sin^2\frac{\theta}{2}\right)^2 = 1 - \cos\theta$. This recovery is optimal as it matches the sums of the squares of the minimal singular values of the initial 4-outcome POVM on $|\phi\rangle$. We note that this probability matches that of initially filtering $|\Psi(\theta)\rangle$ to obtain a maximally entangled state prior to conventional teleportation.

7 Conclusion and Discussion

These results answer a question about the form that measurement trees can take [16]. The general success bound implies that trees with all the final operators conditionally unitary cannot have any non-unitary branch within it. This places strong constraints on the allowed couplings in ADQC-like architectures as all ancilla-driven dynamics much be unitary to maintain the continuing coherence of the register [6]. Even relaxing the requirement for determinism [20], the Cartan decomposition of the system-ancilla interaction must remain rank deficient, i.e. not of the SWAP form [21].

For a binary outcome POVM, the recovery probability takes the form of one minus the difference between the maximum and minimum measurement probabilities. For a multiple outcome POVM, it is the sum of the minimum eigenvalues of each measurement operator, recreating the results of Refs. [8,10] but the proof here is considerably simplified by making no reference to states but emphasizing the spectral structure of the POVM elements. Intuitively, unitarity requires that all states must have the same probability of arriving at the final corrected outcome. After an initial non-ideal measurement, we must cut down transition amplitudes to match that of the least likely state for each branch. This also explains how Procrustean filtering achieves optimality, it selects the largest remaining equal proportion of the non-ideal evolution.

In contrast to Ref. [10], irreversibility is characterized by the ability to discount a state or subspace rather than as a tradeoff against information gain. In quantum state exclusion [22], the task is to minimize the overlap of states with their respective excluding POVM element, minimizing the latter's smallest eigenvalue as a consequence. Hence, *conclusive* state exclusion is irreversible as each outcome has zero overlap with at least one state. This class of measurement is at the opposite end of the spectrum to state discrimination yet is still irreversible.

In the continuous variable limit, a POVM with uncountably many elements all that are proportional to $\mathbb{I} - |\alpha\rangle\langle\alpha|$ (where $|\alpha\rangle$ is some coherent state) [23] reveals an infinitessimal amount of information, but is not reversible for any of its outcomes. This type of measurement produces the maximal gap in the information gain – reversibility tradeoff relation of Eq. 12. For finite dimension d, it is an open question what measurements produce the largest inequality in the tradeoff.

DKLO acknowledges fruitful discussion with John Jeffers, and is supported by Quantum Information Scotland (QUISCO).

References

1. R. Raussendorf, D. E. Browne and H. J. Briegel, Phys. Rev. A **68**, 022312 (2003)
2. E. Kashefi, *et al.*, Electronic Notes in Theoretical Computer Science **249**, 307-331 (2009)
3. J. Anders, *et al.*,Physical Review A **82**, 020301 (2010)
4. J. Anders, *et al.*, Th. Comp. Sci. **430**, 51 (2012)
5. J. Anandan and A. Pines, Phys. Lett. A **141**, 335 (1989)
6. D. K. L. Oi, Phys. Rev. A **89**, 050102(R)(2014)
7. M. Ueda, N. Imoto, and H. Nagaoka, Phys. Rev. A **53**, 3808 (1996)
8. M. Koashi and M. Ueda, Phys. Rev. Lett. **82**, 2598 (1999)
9. G. S. Paraoanu, Phys. Rev. A **83**, 044101 (2011)
10. Y. W. Cheong and S.-W. Lee, Phys. Rev. Lett. **109**, 150402 (2012)
11. A. N. Korotkov, A. N. Jordan, Phys. Rev. Lett. **97**, 166805 (2006)
12. A. N. Jordan, A. N. Korotkov, Contemporary Physics **51**, 125 (2010)
13. G. S. Paraoanu, Euro. Phys. Lett **93**, 64002 (2011)
14. Nadav Katz *et al.*, Phys. Rev. Lett. **101**, 200401 (2008)
15. Yong-Su Kim, Young-Wook Cho, Young-Sik Ra, and Yoon-Ho Kim, Optics Express **17**, 11978 (2009)
16. E. Andersson and D. K. L Oi, Phys. Rev. A **77**, 052104 (2008)
17. J. Dressel, T. A. Brun and A. N. Korotkov, arXiv:1312.1319
18. C. H. Bennett, H. J. Bernstein, S. Popescu and B. Schumacher, Phys. Rev. A **53**, 2046 (1996)
19. P. Agrawal and A. K. Pati, Phys. Lett. A **305**, 12 (2002)
20. K. Halil-Shah, D. K. L. Oi, Proc. TQC 2013, LIPIcs **22**, 1 (2013)
21. J. Zhang, J. Vala, S. Sastry and K. B. Whaley, Phys. Rev. A **67**, 042313 (2003)
22. S. Bandyopadhyay, R. Jain, J. Oppenheim and C. Perry, Phys. Rev. A **89**, 022336 (2014)
23. D. K. L. Oi, V. Potocek and J. Jeffers, Phys. Rev. Lett. **110**, 2010504 (2013)