Transition probabilities and measurement statistics of postselected ensembles

Tobias Fritz[1]

Max Planck Institute for Mathematics[a)]

It is well-known that a quantum measurement can enhance the transition probability between two quantum states. Such a measurement operates after preparation of the initial state and before postselecting for the final state. Here we analyze this kind of scenario in detail and determine which probability distributions on a finite number of outcomes can occur for an intermediate measurement with postselection, for given values of the following two quantities: (i) the transition probability without measurement, (ii) the transition probability with measurement. This is done for both the cases of projective measurements and of generalized measurements. Among other constraints, this quantifies a trade-off between high randomness in a projective measurement and high measurement-modified transition probability. An intermediate projective measurement can enhance a transition probability such that the failure probability decreases by a factor of up to 2, but not by more.

[a)]Electronic mail: fritz@mpim-bonn.mpg.de

1

I. INTRODUCTION

It is a puzzling property of quantum theory that a measurement on a physical system can change the state of that system in a drastic way. A well-known demonstration of this can be made with polarizers (figure 1): upon shining a beam of light onto two orthogonally aligned polarizers, no light at all passes through both of them. However after placing a third polarizer in between the two, such that this new one is not aligned with either of the other two, it is suddenly possible for some light to pass through the whole setup. Hence the middle polarizer, functioning as a projective measurement, has increased the transition probability from zero to a positive value!

Using the simple geometry of a two-state quantum system e.g. in the Bloch sphere, it is not hard to see that the maximal measurement-modified transition probability in this polarizer scenario can be at most $\frac{1}{2}$. But what about other cases like d-dimensional Hilbert spaces of states—what is the maximal modified transition probability then? Or what if the unmodified transition probability does not vanish? And how do the original and the modified transition probability relate to the statistics of the measurement? These questions are what we are concerned with here—mostly for the case of projective measurements, but also for generalized measurements.

The answers to these questions are statements saying that certain things are possible in quantum theory, while other things are not. Hence these answers might in principle be of interest for further high-precision experimental tests of the quantum formalism. For example, it seems conceivable that models with dynamical wavefunction collapse make different predictions than orthodox quantum theory does.

The natural framework for our considerations is the two-state vector formalism of Aharonov and Vaidman[AV07]. There, quantum theory becomes time-symmetric by considering two state vectors for a quantum system: an ordinary state vector $|\psi\rangle$ evolving forward in time, and an additional state vector $|\phi\rangle$ evolving backward in
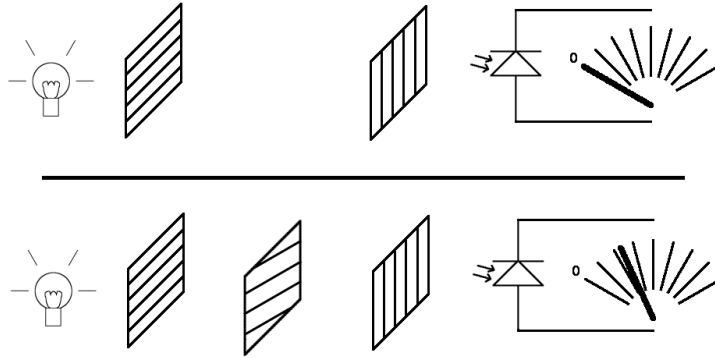
FIG. 1. Enhancing transition probabilities by measurement. We regard the first polarizer as state preparation, the last polarizer as postselection, and the intermediate polarizer as a measurement.

time. While $|\psi\rangle$ is to be interpreted as an initial state, $|\phi\rangle$ should be thought of as a target final state: after all other measurements have been done, the experimenter measures the projection operator $|\phi\rangle\langle\phi|$ and records the results of the experiment in case of a positive result, while discarding the whole run in case of a negative result. This procedure is known as *postselection* with respect to $|\phi\rangle$. The polarizer example from above fits into this framework: the first polarizer can be regarded as preparation of the initial state $|\psi\rangle$, while the final polarizer conducts a postselection with respect to a final state $|\phi\rangle$. The measurement statistics obtained by such a procedure are the statistics of the postselected ensemble. Postselected ensembles can show very counterintuitive behavior: examples are the phenomenon that the so-called "weak value" of an observable can be bigger than the observable's largest eigenvalue[AAV88], or the three-boxes thought experiment which displays a high degree of contextuality[AV07]. Another unexpected property of postselected ensembles has then been found in[Fri10] (see section II), and this is the line of investigation to be continued here. The present article should be readable without knowledge of any of

the references mentioned.

Synopsis. Section II states the problem studied here and recalls some results from a previous paper[Fri10] about dichotomic measurements with postselection. The main result of the present work then follows in section III, a characterization of all triples $(T, S, P(\cdot))$ allowed in quantum theory, where T is the transition probability without measurement, S is the transition probability with measurement, and $P(\cdot)$ is the statistics of the intermediate n-outcome projective (resp. generalized) measurement. After that, section IV discusses some particular special cases of this result and determines to what extent transition probabilities between quantum states can be enhanced by a projective measurement. The latter two sections frequently refer back to the mathematical appendix A. Finally, section V presents a brief conclusion.

Notation. While Dirac notation appears throughout the main text, but not in the mathematical appendix. Sometimes we use expressions like $\langle \phi | A | \psi \rangle$ also when A is not hermitian. In this case, we stipulate that A acts to the right on the vector $|\psi\rangle$.

II. STATEMENT OF THE PROBLEM AND PREVIOUS RESULTS

Outcome probabilities for ensembles with postselection. As has also been outlined in the introduction, we consider a quantum system subject to the following kind of procedure:

(a) prepation of some initial state $|\psi\rangle$,

(b) application of a projective (or generalized) measurement with n outcomes,

(c) postselection$^?$ with respect to some final state $|\phi\rangle$.

We assume that these three consecutive steps happen almost instantaneously, so that the dynamics of the system can be neglected. This is not an essential restriction since we can always take $|\psi\rangle$ to be the actual initial state modified by time evolution until the time of measurement, and similar for $|\phi\rangle$. Also it is no loss of generality to take both $|\psi\rangle$ and $|\phi\rangle$ as pure states, since a mixed state can always be purified by adding an ancilla to the system with which it is entangled (see e.g.$^{\text{NC00}}$ (2.5); for the purification of both $|\psi\rangle$ and $|\phi\rangle$, we might have to add two ancillas).

Concerning the intermediate measurement, we will consider the cases of projective measurement and of generalized measurement separately.

We now calculate the outcome probabilties of the intermediate measurement on the postselected ensemble. The measurement is taken to be defined in terms of Kraus operators V_k, $k \in \{1, \ldots, n\}$, with $\sum_k V_k^\dagger V_k = \mathbb{1}$. It will be assumed for simplicity that the measurement is fine-grained, i.e. that to each outcome k corresponds exactly one Kraus operator V_k; this is enough for our main result III.1, and it should be clear how to extend the following considerations to the general case. With these assumptions, the probability of getting the outcome k in conjunction with successful postselection on the post-measurement state $\dfrac{V_k|\psi\rangle}{\sqrt{\langle\psi|V_k^\dagger V_k|\psi\rangle}}$ is given by the product of

5

the two respective probabilities as

$$\langle\psi|V_k^\dagger V_k|\psi\rangle \cdot \frac{|\langle\phi|V_k|\psi\rangle|^2}{\langle\psi|V_k^\dagger V_k|\psi\rangle} = |\langle\phi|V_k|\psi\rangle|^2.$$

So the cancellation between the normalization of the post-measurement state and the outcome probability gives a surprisingly simple formula for the probability of getting the outcome k in the postselected ensemble:

$$P(k) = \frac{|\langle\phi|V_k|\psi\rangle|^2}{\sum_j |\langle\phi|V_j|\psi\rangle|^2}. \tag{1}$$

Here, the normalization factor

$$S \equiv \sum_j |\langle\phi|V_j|\psi\rangle|^2$$

is the *success probability* of the postselection, i.e. the probability that the final measurement of $|\phi\rangle\langle\phi|$ will give a positive result. We may also regard S as the measurement-modified transition probability. However in order to have a clearer terminology, we will reserve the term "transition probability" for $T = |\langle\phi|\psi\rangle|^2$, and refer to S as the "success probability".

Note that the formalism is time-reversal invariant in the sense that the roles of the initial state and final state can be interchanged without changing the outcome probabilities or the success probability.

For the case of projective measurements, the Kraus operators should be taken to be a complete set of projection operators,

$$V_k = \Pi_k \quad \text{with} \quad \Pi_k^\dagger = \Pi_k, \qquad \Pi_k^2 = \Pi_k, \qquad \sum_k \Pi_k = \mathbb{1}.$$

With this replacement we obtain for the outcome probabilities the Aharonov-Bergmann-Lebowiz formula (eq. (9) in[AV07], see also[ABL64]),

$$P(k) = \frac{|\langle\phi|\Pi_k|\psi\rangle|^2}{\sum_j |\langle\phi|\Pi_j|\psi\rangle|^2}, \tag{2}$$

with the success probability

$$S = \sum_j |\langle \phi | \Pi_j | \psi \rangle|^2$$

as normalization factor.

Introducing the problem. The problem to be solved is the following:

Question II.1. Given the transition probability $T = |\langle \phi | \psi \rangle|^2 \in [0, 1]$, which probability distributions $P(\cdot)$ on $\{1, \ldots, n\}$ are outcome distributions of a projective (resp. generalized) measurement for which values of the success probability $S \in (0, 1]$?

We will only consider the case that the success probability S is strictly positive; for vanishing S, the postselected ensemble is empty, and hence the probability distribution $P(\cdot)$ is not defined.

The operational significance of question II.1 is as follows. All quantities T, S and $P(\cdot)$ are in principle experimentally measurable. We imagine that some experiment has provided us with concrete values for these quantities. Then the task is to find a quantum-mechanical model reproducing these particular values, without specifying the Hilbert space dimension in advance, and assuming that the measurement is projective (resp. generalized). Our main result III.1 then tells us directly whether this is possible or not. Now as already mentioned in the introduction, this could be useful for actual high-precision experimental tests of the quantum formalism, and help to distinguish e.g. models of dynamical wavefunction collapse[Pea99] from orthodox quantum theory, where wavefunction collapse happens instantaneously.

Previous results. The surprising results of[Fri10] have been a strong motivation for the present work. There, the case $T = 0$ and $n = 2$ has been treated in section 2, and it was found that the only possibility is given by $P(1) = P(2) = \frac{1}{2}$, independently of S. This is actually easiest to see on the level of amplitudes, where it follows from

$$0 = \langle \phi | \psi \rangle = \langle \phi | \Pi_1 | \psi \rangle + \langle \phi | \Pi_2 | \psi \rangle,$$

so that the two probabilities for measuring 1 or 2 are given by, respectively,

$$P(1) = \frac{|\langle\phi|\Pi_1|\psi\rangle|^2}{|\langle\phi|\Pi_1|\psi\rangle|^2 + |\langle\phi|\Pi_2|\psi\rangle|^2} = \frac{1}{2}, \qquad P(2) = \frac{|\langle\phi|\Pi_2|\psi\rangle|^2}{|\langle\phi|\Pi_1|\psi\rangle|^2 + |\langle\phi|\Pi_2|\psi\rangle|^2} = \frac{1}{2}.$$

Intuitively, this means that a dichotomic projective measurement with postselection which is orthogonal to the inital state is guaranteed to be a perfectly unbiased random number generator.

III.   MAIN RESULTS

Using the elementary mathematical results listed in appendix A, we are now ready to answer question II.1 in generality.

Theorem III.1.   *(a) A given probability distribution $P(\cdot)$ with given $T \in [0,1]$ and $S \in (0,1]$ can occur via a projective measurement if and only if all the inequalities*

$$\sqrt{P(k)} \le \sqrt{\frac{T}{S}} + \sum_{j \ne k} \sqrt{P(j)} \quad \forall k, \qquad \sqrt{\frac{T}{S}} \le \sum_k \sqrt{P(k)} \le \frac{1}{\sqrt{S}} \qquad (3)$$

hold.

(b) With a generalized measurement, any combination of values for $P(\cdot)$, T and S can occur.

Proof. We start with the proof in the projective measurement case. The main idea here is to use the completeness relation $\sum_k \Pi_k = \mathbb{1}$ in order to obtain an identity for amplitudes

$$\langle\phi|\psi\rangle = \sum_k \langle\phi|\Pi_k|\psi\rangle$$

and then translate this into conditions on the probabilities (2). To this end, we can apply lemma A.1 to

$$z_k = \langle\phi|\Pi_k|\psi\rangle, \;\; k = 1, \ldots, n, \qquad z_{n+1} = -\langle\phi|\psi\rangle.$$

For then upon setting $x_k \equiv \sqrt{P(k)S} = |\langle\phi|\Pi_k|\psi\rangle|$ for $k = 1,\ldots,n$, and defining $x_{n+1} = \sqrt{T}$, it follows that the left-most inequalities of (3) are necessary, as well as the first inequality of the second formula.

The remaining inequality follows from two applications of the Cauchy-Schwarz-inequality as follows:

$$\sum_k |z_k| = \sum_k |\langle\phi|\Pi_k|\psi\rangle| \leq \sum_k \sqrt{\langle\phi|\Pi_k|\phi\rangle}\cdot\sqrt{\langle\psi|\Pi_k|\psi\rangle} \leq \sqrt{\sum_k \langle\phi|\Pi_k|\phi\rangle}\cdot\sqrt{\sum_k \langle\psi|\Pi_k|\psi\rangle} = 1,$$

as was to be shown.

To see that the inequalities (3) taken together are also sufficient for the existence of a quantum-mechanical model, we again set x_k to be given by the square roots of the unnormalized probabilities as $x_k \equiv \sqrt{P(k)S}$ for $k = 1,\ldots,n$, and again define $x_{n+1} = \sqrt{T}$. Then once more by A.1, some compatible z_k's with $\sum_{k=1}^{n+1} z_k = 0$ can now assumed to be given, and they also satisfy $\sum_{k=1}^n |z_k| = \sum_{k=1}^n x_n \leq 1$ by the assumption (3). Now one can use lemma A.2 to obtain the states on $\mathbb{C}^n$ which are given by

$$|\psi\rangle = \sum_{k=1}^n \psi_k|k\rangle, \qquad |\phi\rangle = \sum_{k=1}^n \phi_k|k\rangle$$

in conjunction with the projection operators $\Pi_k = |k\rangle\langle k|$ for $k = 1,\ldots,n$. Then $\sqrt{P(k)S} = |\langle\phi|\Pi_k|\psi\rangle|$ and $T = |\langle\phi|\psi\rangle|^2$ both hold by construction. The requirement $S = \sum_k |\langle\phi|\Pi_k|\psi\rangle|^2$ is automatic by normalization of the probability distribution $P(\cdot)$. This ends the proof in the projective measurement case.

In the generalized measurement case, we will construct $|\psi\rangle$, $|\phi\rangle$ and V_k which reproduce the given data. We first choose any unit vectors $|\psi\rangle$ and $|\phi\rangle$ satisfying $|\langle\phi|\psi\rangle|^2 = T$. Now in Hilbert space of dimension at least n, it is possible to find a complete set of mutually orthogonal projectors Π_k such that $P(k) = \langle\psi|\Pi_k|\psi\rangle$; if one would measure these on the ensemble defined by the initial state $|\psi\rangle$ without postselection, one would obtain the given distribution $P(\cdot)$. Now fix some unit vector

$|\phi'\rangle$ with $|\langle\phi'|\phi\rangle|^2 = S$. Then for those k with $P(k) > 0$, there exists a unitary U_k which maps the unit vector $P(k)^{-1/2}\Pi_k|\psi\rangle$ to $|\phi'\rangle$. We take $V_k \equiv U_k\Pi_k$, while setting $V_k \equiv 0$ for those k with $P(k) = 0$. What we have thus constructed is a generalized measurement in which the post-measurement state is always $|\phi'\rangle$; this guarantees that the measurement statistics on the initial state $|\psi\rangle$ and the postselection are probabilistically independent. Hence by construction, the desired statistics $P(\cdot)$, T and S have been reproduced. $\qquad\square$

It is possible to rewrite the inequalities (3) in a more convenient form. Since the left-most inequality holds for all k if and only if it holds for that k for which $P(k)$ is largest, it is enough to require

$$2\sqrt{\max_k P(k)} \le \sqrt{\frac{T}{S}} + \sum_k \sqrt{P(k)}$$

In terms of the diversity indices (A4), (A5)

$$D_\infty \equiv \frac{1}{\max_k P(k)}, \qquad D_{1/2} \equiv \left(\sum_k \sqrt{P(k)}\right)^2 \tag{4}$$

we can see that the inequalities (3) are in fact equivalent to

$$\boxed{\frac{2}{\sqrt{D_\infty}} - \sqrt{D_{1/2}} \le \sqrt{\frac{T}{S}} \le \sqrt{D_{1/2}} \le \frac{1}{\sqrt{S}}} \tag{5}$$

so that the dependence on the distribution $P(\cdot)$ is only through the dependence on the quantities D_∞ and $D_{1/2}$. By (A6), the allowed interval for $\sqrt{T/S}$ is always non-empty.

Remark III.2. (a) The proof of the theorem shows that it is sufficient to employ Hilbert spaces of dimension at most n. For projective measurements, this is clearly best possible. For generalized measurements however, the number of outcomes is not related to the Hilbert space dimension, and so it might be

10

interesting to study how much the existence of a quantum-mechanical model depends on Hilbert space dimension. Since the proof above still involves many arbitrary choices, it seems conceivable that one can cover a sizeable part of the space of triplets $(P(\cdot), T, S)$ e.g. by qubit models.

(b) The diversity indices D_∞ and $D_{1/2}$ are simply the exponentials of the min-entropy and the Rényi $\frac{1}{2}$-entropy, respectively:

$$H_\infty = \log D_\infty, \qquad H_{1/2} = \log D_{1/2}.$$

(c) The right-most inequality in (5) states that

$$S \le \frac{1}{D_{1/2}}.$$

Intuitively, this means that high randomness in the measurement implies a low success probability. So in order to achieve a high success probability, one needs to choose a projective measurement with not too much randomness on the postselected ensemble.

(d) A very nice example of how to control transition amplitudes by measurements is the Aharonov-Vardi effect[AV80], a variant of the quantum Zeno effect. The observation is that any given quantum dynamics $|\psi(t)\rangle$ can be approximately simulated by starting with the initial state $|\psi(t_0)\rangle$ and conducting projective measurements $|\psi(t_n)\rangle\langle\psi(t_n)|$ at the times $t_n \equiv t_0 + n \cdot \delta t$, with $n \in \mathbb{N}$. Aharonov and Vardi[AV80] have shown in particular that for $\delta t \to 0$, the probability of obtaining any target state $|\psi(t_f)\rangle$ at any final time t_f approaches unity. Since such a sequence of projective measurements can also be seen as a single generalized measurement, this illustrates part (b) of the theorem.

IV. DISCUSSION

Let us now look at some specific cases of theorem III.1(a). So in this section, "measurement" always means "projective measurement".

Case $T = 0$ with S arbitrary. This is the case that has been studied in[Fri10] for $n = 2$. As long as we allow the success probability S to be arbitrarily small, all that remains are the inequalities

$$\sqrt{P(k)} \leq \sum_{j \neq k} \sqrt{P(j)} \quad \forall k \tag{6}$$

For $n = 2$, this reads $\sqrt{P(1)} \leq \sqrt{P(2)}$ and $\sqrt{P(2)} \leq \sqrt{P(1)}$, implying that $P(1) = P(2) = \frac{1}{2}$. Hence a dichotomic measurement with postselection which is orthogonal to the initial state is guaranteed to be a perfectly unbiased random number generator (see section II). The $n = 3$ case is illustrated in figure 2; one obtains a circular disk within the probability simplex. This can be shown from (6) by squaring the inequalities, rearranging, and then squaring again while taking care of the signs. This eventually leads to the quadratic inequalities

$$(P(1) - P(2) - P(3))^2 \leq 2P(2)P(3) \quad + \quad \text{cyclic permutations}$$

for the circular shape of the quantum region in figure 2. Also, just as it should due to the result for the $n = 2$ case, the $n = 3$ region intersects with any side of the triangle in exactly the middle of that side. So whenever the final state is orthogonal to the initial state, any intermediate projective measurement with three outcomes needs to show statistics lying in this disk.

For arbitrary $n \geq 2$, one can at least say that the $\sqrt{P(k)}$ always lie in a certain subset of $\mathbb{R}^n$ which is the convex cone defined by the inequalities (6). Since these n inequalities are linearly independent in $\mathbb{R}^n$, for dimensional reasons this convex cone is a simplex, i.e. the conical hull of n linearly independent extreme rays. One can
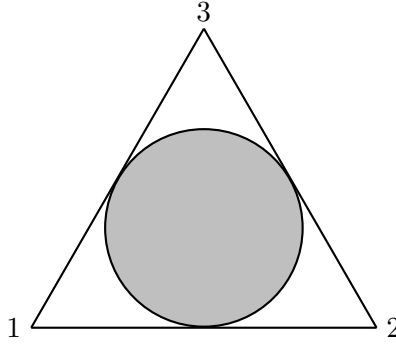
12

FIG. 2. The quantum-mechanical region within the probability simplex for three measurement outcomes, $T = 0$ (orthogonal postselection), and arbitrary success probability S. This is a ternary plot, i.e. each vertex stands for a definite outcome, and each point inside the triangle represents a probability distribution over the vertices.

calculate the mth extreme ray by requiring all inequalities except for the mth one to be saturated. Solving the ensuing system of linear equations shows that the mth extreme ray y^m has the coordinates

$$y_j^m = 1 + (2 - n)\delta_{jm}$$

Hence for any $P(\cdot)$ satisfying (6), one can find non-negative real numbers λ_m such that

$$\sqrt{P(k)} = \sum_m \lambda_m y_k^m.$$

Case S fixed, P(·) fixed. The first two inequalities of (5) define an interval of possible values for the transition probability T. This can be interpreted as follows: by knowing the behavior of the system with measurement, it is possible to predict something about how the system would behave without measurement.

Case T > 0 fixed, P(·) fixed, S arbitrary. Here, it is possible for any $P(\cdot)$ to find some appropriately small success probability S such that all inequalities in (5) hold
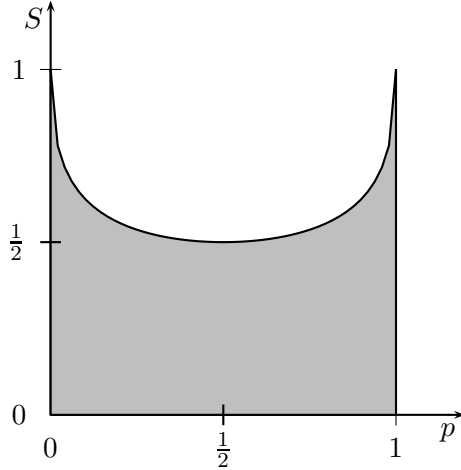
FIG. 3. For $n = 2$ (dichotomic measurement), the possible quantum-mechanical success probabilities S as a function of the outcome probability p. High randomness in the measurement decreases the maximal probability of successful postselection, i.e. the maximal measurement-modified transition probability.

(e.g. $S = \frac{T}{D_{1/2}}$), so no constraints abound. This is one reason why it is important to always consider S as an additional parameter.

Case $n = 2$ with T and S unspecified. Here, the two probability values $P(1)$ and $P(2)$ determine each other uniquely, so let us write $P(1) = p$ and $P(2) = 1 - p$. Then the inequalities are

$$\left| \sqrt{p} - \sqrt{1-p} \right| \leq \sqrt{\frac{T}{S}} \leq \sqrt{p} + \sqrt{1-p} \leq \frac{1}{\sqrt{S}} \tag{7}$$

The projection of this into the p-S-plane, where only the last inequality is relevant, is shown in figure 3. For fixed S, some sections of the quantum region are graphed in figure 4. The first two inequalities of (7) define the upper and lower boundary curves in these figures, while the third inequality leads to vertical cuts whenever $S > \frac{1}{2}$.
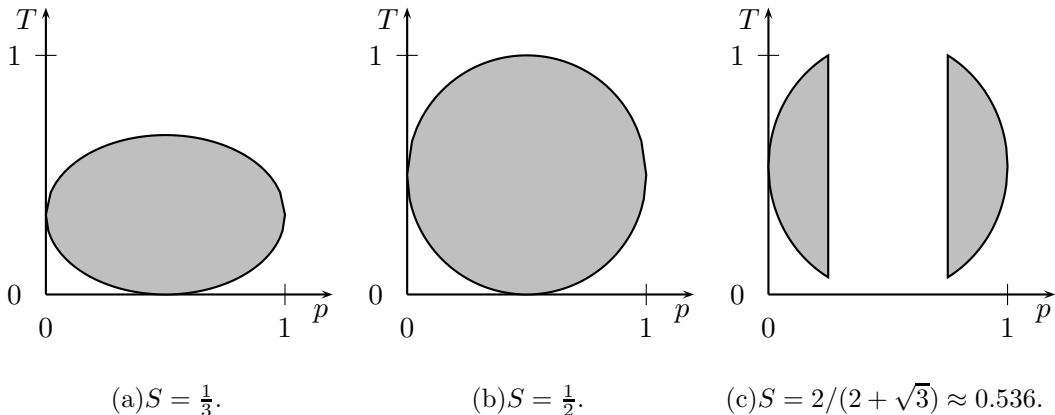
14

(a)$S = \frac{1}{3}$.   (b)$S = \frac{1}{2}$.   (c)$S = 2/(2 + \sqrt{3}) \approx 0.536$.

FIG. 4. Again $n = 2$ (dichotomic measurement). These plots show the quantum-mechanical region for (p, T) for some values of S. The vertical cuts for $S > \frac{1}{2}$ appear due to figure 3. For $S \to 1$, these cuts rapidly approach the $p = 0$ and $p = 1$ axes. One possible interpretation is that knowing the system behavior with measurement (i.e. S and p) lets us say something about system behavior without measurement (i.e. T).

The T-S-region. How does the transition probability relate in general to the probability of successful postselection? To study this, it is best to consider the inequalities in the form (5). Figure 5 shows an illustration of the following proposition.

Proposition IV.1. *For a given number of outcomes n, some success probability S can appear in quantum theory together with some transition probability T if and only if*

$$\boxed{\frac{T}{n} \leq S \leq \frac{T + 1}{2}} \tag{8}$$

Proof. Again it is first shown that these inequalities are necessary. Since $D_{1/2} \leq n$, the second inequality in (5) implies that

$$T \leq nS.$$

For proving the second inequality of (8), we distinguish two cases. If, firstly, the left-most term of (5) is non-negative, we can square the left-most inequality of (5)

15

and use it as follows:

$$\frac{T}{S} + \frac{1}{S} \overset{(5)}{\geq} \frac{4}{D_\infty} + 2D_{1/2} - 4\sqrt{\frac{D_{1/2}}{D_\infty}} = 2\left[\frac{1}{D_\infty} + \left(\sqrt{D_{1/2}} - \frac{1}{\sqrt{D_\infty}}\right)^2\right],$$

so that the desired result follows from lemma A.4(a). The second case is that the left-most term of (5) is negative, which means that $D_{1/2} \geq \frac{4}{D_\infty}$. If $D_\infty \leq 2$, we have therefore $D_{1/2} \geq 2$, so that $S \leq \frac{1}{2}$ by $S \leq \frac{1}{D_{1/2}}$. Finally if $D_\infty \geq 2$, then (A6) also shows that $D_{1/2} \geq 2$, giving the same conclusion $S \leq \frac{1}{2}$. In all cases, the second inequality of (8) has therefore been verified.

For checking sufficiency of (8), consider first the case that $\frac{T}{n} \leq S \leq T$. Then by lemma A.4(b), it follows that the first inequality of (5) holds automatically. The possible values for $D_{1/2}$ are given by the closed interval $[1, n]$. Hence it is possible to find some value for $D_{1/2}$ in this interval which also satisfies (5) whenever $\frac{1}{\sqrt{S}} \geq 1$, which holds trivially, and $\sqrt{\frac{T}{S}} \leq \sqrt{n}$, which is true by assumption. This ends the proof in this case.

It remains to prove sufficiency when $T \leq S \leq \frac{T+1}{2}$. Here, it is in fact enough to consider probability distributions $P(\cdot)$ supported on two elements, which brings us effectively down to the dichotomic case $n = 2$ from equation (7). By $\sqrt{\frac{T}{S}} \leq 1$, the middle inequality of (7) is automatic, so one only needs to take care of the remaining two. These in turn can be written as

$$1 - 2\sqrt{p(1-p)} \leq \frac{T}{S}, \qquad 1 + 2\sqrt{p(1-p)} \leq \frac{1}{S}$$

Upon choosing p such that the second inequality is saturated, one finds that the first inequality is satisfied as long as $\frac{T+1}{S} \geq 2$. $\qquad\square$

So this result gives clear bounds on how much a measurement can enhance or reduce a transition probability. It has been found that a measurement can reduce a transition probability by a factor which is given by the number of outcomes of the

16

measurement. This becomes intuitive when one thinks of the measurement—with outcomes discarded—as a decoherence process which can drive the system's state towards a totally mixed state or highly mixed state.

The situation for enhancing transition probabilities by measurement is very different. We can rewrite the first inequality of (8) more conveniently in terms of the *failure probabilities* $1 - T$ and $1 - S$, where it reads

$$\boxed{1 - S \geq \frac{1 - T}{2}}$$

Hence, a measurement can lower the probability that a desired state transition fails by a factor of up to 2, but not by more. The proof above has shown that this enhancement can already be achieved by a two-outcome measurement. This is again intuitive in terms of the decoherence due to measurement: for creating a successful transition, it would be useless to try to measure a projection operator with support outside of the linear span $\mathrm{lin}\,\{|\psi\rangle, |\phi\rangle\}$. Therefore, a transition-enhancing measurement should have non-vanishing probability on exactly two outcomes.

V.   CONCLUSION

It is a well-known phenomenon that measurements influence transition probabilities between quantum states. In this article, we have conducted a systematic study of this phenomenon and determined how it relates to the outcome distribution of the intermediate measurement on the corresponding postselected ensemble. It has been found that a given probability distribution for a projective measurement can appear in conjunction with a given transition probability and a given success probability of the postselection if and only if certain inequalities hold. These inequalities depend on the probability distribution only through its min-entropy and its Rényi $\frac{1}{2}$-entropy. Furthermore, no conditions at all abound if the measurement is allowed to be any
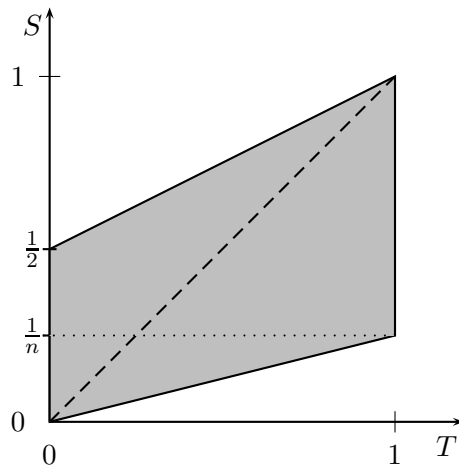
FIG. 5. The quantum region of transition probabilities: T is the transition probability without measurement, while S is the transition probability with n-ary projective measurement. All points above the dashed diagonal $S = T$ represent a measurement-enhanced transition probability.

generalized quantum measurement.

As a consequence of these results, it was possible to bound the enhancement of transition probabilities by projective measurements. The maximal enhancement can be achieved with two-outcome measurements and is such that the failure probability decreases by a factor of 2.

REFERENCES

[AAV88] Yakir Aharonov, David Z. Albert, and Lev Vaidman. How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100. *Phys. Rev. Lett.*, 60(14):1351–1354, Apr 1988.

[ABL64] Yakir Aharonov, Peter G. Bergmann, and Joel L. Lebowitz. Time symmetry

in the quantum process of measurement. *Phys. Rev.*, 134(6B):B1410–B1416, Jun 1964.

[AV80]Y. Aharonov and M. Vardi. Meaning of an individual "Feynman path". *Phys. Rev. D*, 21(8):2235–2240, Apr 1980.

[AV07]Yakir Aharonov and Lev Vaidman. The two-state vector formalism: An updated review. In *Time in Quantum Mechanics*, volume 734 of *Lecture Notes in Physics*, pages 399–447. Springer Berlin / Heidelberg, 2007.

[Fri10]Tobias Fritz. On the existence of quantum representations for two dichotomic measurements. *Journal of Mathematical Physics*, 51(5):052103, 2010.

[HSS03]A. Higuchi, A. Sudbery, and J. Szulc. One-qubit reduced states of a pure many-qubit state: Polygon inequalities. *Phys. Rev. Lett.*, 90(10):107902, Mar 2003.

[Jos06]Lou Jost. Entropy and diversity. *Oikos*, 113(2):363–375, 2006.

[NC00]Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information.* Cambridge University Press, Cambridge, 2000.

[Pea99]Philip Pearle. Collapse models. In *Open Systems and Measurement in Relativistic Quantum Theory*, volume 526 of *Lecture Notes in Physics*, pages 195–234. Springer Berlin / Heidelberg, 1999.

[Pin05]Iosif Pinelis. Cyclic polygons with given edge lengths: existence and uniqueness. *J. Geom.*, 82(1-2):156–171, 2005.

Appendix A: Mathematical appendix

Here we collect various elementary mathematical facts which are referenced from the main text.

It is known (see e.g.[Pin05]) that a finite sequence of non-negative real numbers, $x_1, \ldots, x_n$ is the sequence of edge lengths of a polygon in the Euclidean plane if and

only if the inequalities

$$x_k \leq \sum_{j \neq k} x_j \tag{A1}$$

hold. Since they generalize the triangle inequality, these inequalities are known as *polygon inequalities*; on the other hand, the triangle inequality immediately implies that the polygon inequalities are necessary for the existence of a polygon with these edge lengths. See e.g.[HSS03] for another occurence of the polygon inequalities in quantum information theory.

This geometrical statement directly implies the following:

Lemma A.1. *Given non-negative real numbers $x_1, \ldots, x_n$, there exist complex numbers $z_1, \ldots, z_n$ with*

$$|z_k| = x_k, \qquad \sum_k z_k = 0$$

if and only if the inequalities

$$x_k \leq \sum_{j \neq k} x_j \tag{A2}$$

hold.

Lemma A.2. *For $n \geq 2$ and any $z \in \mathbb{C}^n$, there exist $\psi, \phi \in \mathbb{C}^n$ with*

$$||\psi||_2 = 1 = ||\phi||_2, \qquad \overline{\psi}_k \phi_k = z_k \quad \forall k = 1, \ldots, n$$

if and only if the inequality

$$\sum_k |z_k| \leq 1 \tag{A3}$$

holds.

Proof. Necessity of (A3) is nothing but the Cauchy-Schwarz inequality:

$$\sum_k |z_k| = \sum_k |\psi_k| \cdot |\phi_k| \leq \sqrt{\sum_k |\psi_k|^2} \cdot \sqrt{\sum_k |\phi_k|^2} \leq 1$$

That (A3) is also sufficient for the existence of such ψ and ϕ will be shown by induction on n. Note that the phases of z_k can be changed arbitrarily without altering the (non-)existence of such vectors, hence we may as well assume that all z_k are non-negative real numbers. We now prove the statement for the initial case $n = 2$. By the assumptions $z_1, z_2 \geq 0$ and $z_1 + z_2 \leq 1$, it is implied that $|z_1 - z_2| \leq 1$, and therefore it is possible to find angles α and β such that

$$\cos(\alpha + \beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta \overset{!}{=} z_1 - z_2,$$
$$\cos(\alpha - \beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta \overset{!}{=} z_1 + z_2.$$

Hence the two vectors

$$\psi = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}, \qquad \phi = \begin{pmatrix} \cos\beta \\ \sin\beta \end{pmatrix}$$

have all the required properties.

The induction step is a simple rescaling argument. Given $z_1, \ldots, z_{n+1} \geq 0$ with $\sum_k z_k \leq 1$, define $z'_1, \ldots, z'_n$ as

$$z'_k \equiv \frac{z_k}{1 - z_{n+1}}, \quad k = 1, \ldots, n.$$

(We may assume $z_{n+1} \neq 1$ e.g. by reordering the z_k's.) Then by induction assumption, we can find $\psi', \phi' \in \mathbb{C}^n$ with $||\psi'||_2 = ||\phi'||_2 = 1$ and $\overline{\psi'_k}\phi'_k = z'_k$. Now the two vectors

$$\psi_k \equiv \begin{cases} \psi'_k\sqrt{1 - z_{n+1}} & \text{for } k = 1, \ldots, n \\ \sqrt{z_{n+1}} & \text{for } k = n + 1 \end{cases}, \qquad \phi_k \equiv \begin{cases} \phi'_k\sqrt{1 - z_{n+1}} & \text{for } k = 1, \ldots, n \\ \sqrt{z_{n+1}} & \text{for } k = n + 1 \end{cases},$$

do indeed have the desired properties $||\psi||_2 = ||\phi||_2 = 1$ and $\overline{\psi_k}\phi_k = z_k$, which also finishes the induction step. $\qquad\square$

The following fact can also be regarded as a special case of the Hölder inequality, but since a direct proof is extremely simple, we have included it here.

Lemma A.3. *Let $x \in \mathbb{R}_{\geq 0}^n$ with $\sum_k x_k = 1$. Then,*

$$\left(\sum_k \sqrt{x_k} \right)^2 \cdot \max_k x_k \geq 1$$

Proof. This is easily shown by a direct calculation:

$$1 = \sum_k \sqrt{x_k} \cdot \sqrt{x_k} \leq \left(\sum_k \sqrt{x_k} \right) \cdot \max_k \sqrt{x_k}$$

so that squaring gives the desired result. $\qquad\square$

Diversity indices. A diversity index[Jos06], as used for example in biostatistics, is a function that assigns to each probability distribution a real number which is intended to measure a sort of effective cardinality contained in the probability distribution. In other words, a diversity index is an exponentiated entropy. Like in the main text, let $P(\cdot)$ be a probability distribution on $\{1, \ldots, n\}$. For each $q \in (0, \infty)$, one obtains a diversity index D_q by defining

$$D_q(P) \equiv \left(\sum_k P(k)^q \right)^{\frac{1}{1-q}} \tag{A4}$$

For $q = 1$, this has to be understood as $\lim_{q \to 1} D_q$, which is the exponentiated Shannon entropy. In a similar way, it is possible to define $D_0(P)$, which is the cardinality of the support of P, and $D_\infty(P)$, which turns out to be

$$D_\infty(P) = \frac{1}{\max_k P(k)}. \tag{A5}$$

The relevant quantities for us are going to be $D_{1/2}$ and D_∞. When P is the uniform distribution on n elements, we have $D_q(P) = n$ for all q.

In this notation, we get a simple reformulation of lemma A.3:

$$D_\infty \leq D_{1/2}. \tag{A6}$$

Lemma A.4. *(a)*

$$\frac{1}{D_\infty} + \left(\sqrt{D_{1/2}} - \frac{1}{\sqrt{D_\infty}}\right)^2 \geq 1$$

(b)

$$\frac{2}{\sqrt{D_\infty}} - \sqrt{D_{1/2}} \leq 1$$

Proof. (a) Let $k_0 \in \{1, \ldots, n\}$ be such $P(k_0)$ is the highest probability in the distribution, i.e. $P(k_0) = \max_k P(k)$. Then,

$$\frac{1}{D_\infty} + \left(\sqrt{D_{1/2}} - \frac{1}{\sqrt{D_\infty}}\right)^2 = P(k_0) + \left(\sum_{k \neq k_0} \sqrt{P(k)}\right)^2 \geq P(n) + \sum_{k \neq k_0} P(k) = 1,$$

(b) This is trivial by $D_\infty \geq 1$ and $D_{1/2} \geq 1$.

$\square$

23