Multitime quantum communication: Interesting but not counterfactual

Robert B. Griffiths ⓘ*

Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA

A protocol for transmission of information between two parties introduced by Salih *et al.* [Phys. Rev. Lett. **110**, 170502 (2013)] involves sending quantum amplitude back and forth through a quantum channel in a series of steps, rather than simply sending a signal in one direction. The authors claimed that their protocol was "counterfactual" in the sense that while a quantum channel is needed to connect the parties, its actual usage becomes vanishingly small in the asymptotic limit as the number of steps tends to infinity. Here we show that this claim is incorrect because it uses probabilistic reasoning that is not valid at intermediate times in the presence of quantum interference. When ill-defined probabilities are replaced with a well-defined measure of channel usage, here called "Cost," equal to the absolute square of the amplitude sent through the channel, the total Cost does not go to zero in the asymptotic limit of a large number of steps, but is bounded below by a rigorous inequality. A detailed analysis shows that this bound is satisfied in the protocol of Salih *et al.* [Phys. Rev. Lett. **110**, 170502 (2013)]. The analysis leading to the bound uses the fact that the Gram matrix formed by inner products of a collection of pure quantum states is additive over Hilbert subspaces and invariant under unitary time transformations. Its off-diagonal elements, which in general are not positive, play a significant role in the formal argument as well as providing a somewhat strange way of visualizing the transfer of information.

I. INTRODUCTION

The motivation for this paper is a scheme for the transmission of quantum information introduced by Salih *et al.* [1] and referred to hereafter as SLAZ, the initials of the authors. One ordinarily thinks of the transmission of information as sending a signal through a channel from a sender to a receiver. However, the idea in SLAZ is that information can be sent from Bob to Alice if the quantum particle used to carry the information starts off in Alice's domain and a part of its quantum amplitude is sent to Bob through a quantum channel. Bob modifies this is some way before sending (or possibly not sending) it back to Alice, depending on the signal he wants to send. Alice then employs what Bob has returned to begin a second round of sending amplitude to Bob, who again modifies it before returning it, and so forth. This back-and-forth motion can continue for a large number of rounds until the information that Bob is sending has arrived in Alice's domain, where she can carry out a measurement or perhaps perform additional processing. A key feature of protocols of this type is that all the intermediate steps can be represented by purely unitary time evolution, with intermediate time measurements, if any, replaced by unitaries, a process of purification.

The use of *amplitude* rather than *particle* in the preceding paragraph is intentional, because the state of the photon or other particle is in general a coherent superposition of parts associated with different spatial locations: Alice's domain, Bob's domain, and the channel connecting them. One generally thinks of a particle as something with a spatial location, but in quantum mechanics one cannot simultaneously ascribe particle and wave properties to the same entity at the same

time because of wave-particle duality. In Hilbert-space quantum mechanics, physical properties, such as location in space, are represented by projectors, see Sec. II.5 of [2], and when a projector representing a wave, think of $|\psi\rangle\langle\psi|$, does not commute with a projector specifying a spatial location, ignoring this fact can rapidly lead to paradoxes. The double-slit paradox is an example: When a coherent wave passes through the slit system one cannot say through which slit the particle passed.

The term "counterfactual" in the original SLAZ paper has the following significance. A quantum channel connecting the communicating parties is essential: This is not a case of mysterious nonlocal influences of the sort which are sometimes invoked to explain quantum violations of Bell inequalities. However, if the number of steps in an SLAZ protocol is sufficiently large, the magnitude of the amplitude sent through the channel in each step can be made very small and vanishes in the limit as the number of steps tends to infinity.

A similar claim of counterfactuality has been made in much of the rather substantial literature motivated by the original SLAZ publication, which contains various modifications and extensions of the original protocol. There have also been criticisms of these counterfactual claims and (of course) replies to criticisms. The Conclusion (Sec. V) of the present paper contains a few remarks about how its results apply to some of these publications, but a review, much less a detailed discussion, lies far outside its scope. The interested reader is referred to the extensive bibliographies found in [3,4].

The aim of the present paper is to study the use of quantum channels in protocols of the SLAZ type, in particular the sense in which this usage is or is not counterfactual. To this end a technical term, *Cost*, the absolute square of the amplitude passing through the channel in a particular step in the pro-

tocol, is used for reasons discussed in Sec. II, as a useful substitute for "probability", which in a quantum context is often ill-defined. The example of multiple channels in parallel, which few would want to claim are counterfactual, serves as an introduction to how information can be sent through a single channel in a single direction at multiple times, in a process in which all of the intermediate steps are represented by unitary maps.

The main mathematical results of this paper are in Sec. III: Gram matrices and some of their properties are discussed in Sec. III A, while Sec. III B gives the basic structure of simple two-way multiple time protocols. Section III C considers simple schemes for transmitting one classical bit, while the rigorous lower bound that undermines various counterfactual claims is the topic of Sec. III D.

The original SLAZ protocol is studied in detail in Sec. IV. In particular, the total Costs of transmitting a classical bit $\lambda = 0$, in which Bob reflects the amplitude back to Alice, and for transmitting $\lambda = 1$, in which he absorbs rather than returns it, are evaluated explicitly. It turns out that in the asymptotic limit the $\lambda = 1$ Cost is minuscule, but that for $\lambda = 0$ is enormous, while the product of the two remains finite and satisfies the rigorous bound in Sec. III D. The mistaken claim that the SLAZ protocol is counterfactual results from two errors: a concept of channel use which would be questionable even for a classical stochastic process and an improper use of probabilities in a way that violates quantum principles.

Section V has a summary of the main results of this paper, a few comments on some parts of the literature related to SLAZ, and some suggestions for future directions of research. This author believes that protocols of the SLAZ type are quite interesting, deserve further exploration, and might contribute to useful ways of studying multipartite and multitime transmission of quantum information, as in quantum networks. And that such studies would prove more fruitful in the absence of claims of counterfactuality.

II. ONE-WAY PROTOCOLS

A. Multiple channels in parallel

Think of quantum information as the information carried by a photon as it passes through a quantum channel, such as an optical fiber. The information could be encoded in its polarization. Rather than using a single channel, one could imagine sending the photon as a superposition state through a set of N channels in parallel, using a collection of beam splitters to divide up the initial amplitude among the different channels, and a corresponding collection to later recombine them. Let us suppose that the normalized $|\Phi\rangle$ that represents the photon at some intermediate time is a coherent superposition of amplitudes

$$|\Phi\rangle = \sum_{n=1}^{N} c_n |\phi_n\rangle \tag{1}$$

associated with the individual channels, labeled by n. Define the Cost q_n associated with the use of channel n and the total

Cost Q for the channel system as

$$q_n := |c_n|^2, \quad Q := \sum_{n=1}^{N} q_n. \tag{2}$$

If the $|\phi_n\rangle$ and $|\Phi\rangle$ are normalized, Q is equal to 1, so one might identify q_n with the probability that the photon is in channel n. But what does that mean? In standard (textbook) quantum mechanics probability refers to the outcome of a measurement, but a measurement carried out at an intermediate time, when the quantum state is a coherent superposition over various locations, can alter what occurs later, and hence it is dangerous to associate such a probability with a situation in which a measurement does *not* take place.

Another way of viewing this difficulty is to recall that von Neumann (Sec. III.5 of [2]) identified quantum *physical properties*—which in classical physics are associated with sets of points in the classical phase space—with projectors, self-adjoint idempotent operators $P = P^\dagger = P^2$ on the quantum Hilbert space. For example, in the case of a spin-half particle the projectors

$$P = (I - \sigma_z)/2, \quad R = (I + \sigma_x)/2, \tag{3}$$

where I is the identity and σ_z and σ_x are Pauli operators, represent the properties $S_z = -\hbar/2$ and $S_x = +\hbar/2$, respectively. In general, if two projectors P and R commute their product $PR = RP$ represents the property P AND R. However, if they do not commute, neither PR nor RP is a projector, and so neither represents a quantum property. In some sense noncommutation is the very essence of quantum mechanics; it is what distinguishes it from classical physics. The use of standard (Kolmogorov) probabilities requires a *sample space* of mutually exclusive possibilities, one and only one of which occurs in a particular run of an experiment. In quantum theory such a sample space is a collection of mutually orthogonal projectors that sum to the identity, a *projective decomposition of the identity*. For example, R and $I - R$ in (3) in the case of spin half; see (7) below for the general definition. In quantum mechanics there are often many possible sample spaces that one might be interested in, and carelessly combining incompatible spaces—some projectors in one do not commute with projectors in the other—inevitably leads to paradoxes rather than physical understanding.

In the present context the dyad $|\Phi\rangle\langle\Phi|$ is a projector that does not commute with any of the projectors $|\phi_n\rangle\langle\phi_n|$ for which c_n is nonzero, and thus it is meaningless to assign a probability to the latter in a situation where the coherent superposition $|\Phi\rangle$ will later be transformed by the final beam splitters into the original state that entered the channel system. For example, in a double-slit experiment in which the amplitudes from the two slits combine coherently to produce interference, it makes no sense to talk about the probability that the photon previously passed through one slit rather than the other. The two Costs are well defined: They are simply the absolute squares of the two amplitudes. However, attempting to measure which slit the particle passes through in order to define a probability will destroy the interference pattern.

B. One channel used multiple times

The possible advantages, if any, of using many channels in parallel can also be realized by employing a single channel and sending quantum amplitude through it at a succession of times; this is what makes protocols of the SLAZ type of some interest. Let us suppose that information is being sent from Bob to Alice. One can think of the photon at a particular time as being in a coherent superposition of amplitudes in three different physical locations: Alice's domain A, Bob's domain B, and the channel C connecting them. The same symbols can be used for the Hilbert-space projectors associated with these locations, thus operators which are self-adjoint and idempotent, $A = A^\dagger = A^2$, and mutually orthogonal, $AB = BC = AC = 0$. They sum to the identity

$$A + C + B = I \tag{4}$$

and hence form a projective decomposition of the identity [see the general definition in (7) below]. A projective decomposition of the identity is the quantum counterpart of the sample space of mutually exclusive possibilities essential for using standard (Kolmogorov) probability theory in the case of a quantum system. Note that A, B, and C are *subspaces* of a single Hilbert space, not *subsystems* represented by a tensor product. If the quantum particle possesses other degrees of freedom, these projectors are to be understood using the usual convention as including the identity operator on these additional degrees of freedom. Thus for a photon, A means that it is located in Alice's domain, whatever may be its polarization.

Bob can send a particular type of information λ to Alice by starting with a normalized reference state $|\psi_0\rangle = B|\psi_0\rangle$, the particle is somewhere in his domain B, and using a unitary transformation $\mathcal{B}^\lambda$ acting on the subspace $B + C$ to place it in a state

$$\left|\psi_1^\lambda\right\rangle = C\left|\psi_1^\lambda\right\rangle = \mathcal{B}^\lambda|\psi_0\rangle \tag{5}$$

in the channel, at which point it travels through the channel to Alice. As the channel has no effect except to transmit the particle from one end to the other, we simplify the discussion (here and later) by using the same symbol for the ket that arrives at Alice's end. She then applies a unitary $\mathcal{A}$ that does not depend on λ, for she does not know what Bob is sending, to empty the channel and arrive at a final state

$$\left|\psi_2^\lambda\right\rangle = A\left|\psi_2^\lambda\right\rangle = \mathcal{A}\left|\psi_1^\lambda\right\rangle, \tag{6}$$

which she can then measure or subject to further processing.

This single-round transmission process can be carried out in a number of rounds in which during the nth round Bob employs a unitary $\mathcal{B}_n^\lambda$ acting on the $B + C$ subspace to map an amplitude $c_n|\psi_0\rangle$ into C, which is initially empty, and which travels to Alice, who uses a unitary $\mathcal{A}_n$ acting on $A + C$ to remove it from the channel, which is then empty and ready for the next round. One way to visualize this is that Bob has a domain B of high dimension and at the outset splits up the initial amplitude $|\psi_0\rangle$ into pieces placed in different subspaces of B with the help of a suitable set of beam splitters. At round n the unitary $\mathcal{B}_n^\lambda$ interchanges the appropriate subspace of B with the empty C. Alice's A is also large and her $\mathcal{A}_n$ maps whatever Bob has sent into an empty subspace reserved for this purpose. When the run is completed Alice can then combine the

amplitudes in these different subspaces into a smaller space, e.g., using beam splitters, or she can do a similar combination at the end of each round. Of course Alice's and Bob's unitaries cannot be chosen independently; the two must work together to design the protocol. What is unknown to Alice is Bob's choice of λ for a particular run; this is the information that she can extract at the end.

Some multiple-time protocols employ measurements by Alice at intermediate times. In cases such as the original SLAZ scheme, discussed below in Sec. IV, it is possible to store the amplitude that could have triggered the measuring device in an empty subspace in Alice's domain and put off the measurement until the protocol is finished. Of course, amplitudes that correspond to several measurements in succession can be combined, just as in the case of simultaneous transmission through several channels in parallel, as discussed earlier.

III. TWO-WAY PROTOCOLS

A. Gram matrices

Let $\{P_j\}$ be a projective decomposition of the Hilbert space identity I,

$$I = \sum_j P_j, \quad P_j = P_j^\dagger, \quad P_j P_k = \delta_{jk} P_j, \tag{7}$$

and let $\{|\psi^\mu\rangle\}$, $\mu = 0, 1, \ldots$, be a collection of kets on the same Hilbert space. The *Gram matrix*

$$G^{\mu\nu} = \langle\psi^\mu|\psi^\nu\rangle = \sum_j G^{\mu\nu}(P_j) = \sum_j \langle\psi^\mu|P_j|\psi^\nu\rangle \tag{8}$$

is *additive* in that it is a sum over contributions from the different subspaces. In addition, $G^{\mu\nu}$ is *invariant* (or *conserved*) under a unitary operation U that acts on every ket in the collection $\{|\psi^\mu\rangle\}$. Also, if this unitary acts on only some of the subspaces, say, P_1 and P_2, and is the identity operator on the others, then while both $G^{\mu\nu}(P_1)$ and $G^{\mu\nu}(P_2)$ may change, their sum $G^{\mu\nu}(P_1) + G^{\mu\nu}(P_2)$ remains unchanged. That Gram matrices are additive and conserved plays an important role in what follows.

We will refer to the diagonal elements $G^{\mu\mu}(P_j)$, which are non-negative, as *weights*. As these are rather like probabilities, their additivity and conservation is not surprising. However, that the same is true of the nondiagonal elements $G^{\mu\nu}(P_j)$ with $\mu \neq \nu$, hereafter referred to as *overlaps*, comes as something of a surprise, especially since $|\psi^\mu\rangle$ and $|\psi^\nu\rangle$ may refer to two different runs of an experiment, one on Friday and one on Monday. Nonetheless, overlaps play a key role in the following analysis, not only as part of the mathematics but also in a surprising but useful intuitive way of thinking about what is going on. The absolute value of an overlap corresponds to a notion of *fidelity* in quantum information, but in general an overlap is a complex number, and the fact that it can be negative as well as positive is a key element in what follows.

B. Basic two-way protocol

In the following discussion the projective decomposition of the identity (7) that will concern us is $\{A, C, B\}$, where A means that the photon or other quantum particle is in Alice's

domain, B that it is Bob's domain, and C that it is in the channel connecting them. At the beginning of a two-way protocol of the SLAZ type, in which Bob is sending information to Alice, all of the photon amplitude is in Alice's domain A. She initiates the run by sending some amplitude to Bob through the channel. He then modifies it and returns some or all of it to Alice, in a manner that depends on the information λ he wishes to transmit. Alice processes what Bob has returned and begins the second round by again sending amplitude to Bob, who again returns it, etc. This can go on for N rounds, following which Alice makes a measurement to determine the value of λ.

In further detail: At the beginning of round n, Alice uses a unitary $\mathcal{A}_{n1}$ acting on $A + C$ to map some of the amplitude in A into an empty channel C. This amplitude then flows through the channel to Bob, where he empties the channel into B, does some processing, and then maps some amplitude back into C. This flows to Alice, who empties C into A using a unitary $\mathcal{A}_{n2}$. We assume that "flow through the channel" does not change anything, and hence it is convenient not to think of C as divided into close-to-Alice, close-to-Bob, and in-between subspaces, but simply imagine that Alice and then Bob and then Alice are acting on a single C subspace. Alice uses unitaries that act on $A + C$ and are independent of λ, while Bob uses unitaries $\mathcal{B}_n^\lambda$, that depend on the information λ he wants to transmit, which act on $C + B$. Both the Alice and Bob unitaries will in general depend upon the round n, but Alice's do not depend upon λ. In addition, we impose the restriction that Bob's actions are passive in the sense that the magnitude of the amplitude he sends back to Alice in round n cannot be greater than what he has just received. This last condition clearly differentiates these two-way protocols from the one-way protocols of Sec. II B.

The requirement that Alice and Bob only employ unitary operations simplifies the analysis. It is true that various published protocols of this type, including the original SLAZ version to be discussed in Sec. IV, employ nonunitary measurements at intermediate times. In the cases of interest to us these measurements can be replaced by unitary operations which allow the measurements to be put off until the end of the run, in a manner indicated in Sec. II B and employed in the discussion in Sec. IV.

To quantify the channel usage for these protocols we use the notions of Cost, equal to the absolute square of the amplitude for a single use of the channel, and Total Cost for the sum of the Costs involved in a single experimental run, as in Sec. II A [see (2)]. An important issue connected to claims that these protocols are counterfactual has to do with the difference between Cost and probability, as will be discussed later for the SLAZ protocol in Sec. IV (the importance of this has already been noted in Sec. II A). In particular, we will be interested in identifying protocols that minimize the overall Cost, as in the example discussed next.

C. Sending one classical bit

In the simplest SLAZ protocol Bob wants to send a single classical bit $\lambda = 0$ or 1 to Alice. At the start all of the amplitude is in A for both a $\lambda = 0$ and a $\lambda = 1$ run, so all four of the initial Gram matrix elements $G_0^{\mu\nu}(A)$, where μ and ν are

the possible values of λ, are equal to 1. The goal is that after N rounds the Gram matrix will be

$$G_N^{\mu\nu}(A) = \delta_{\mu\nu}, \qquad (9)$$

that is to say, the final result for a $\lambda = 0$ run is orthogonal to that for a $\lambda = 1$ run, and Alice, by making a measurement in an appropriate basis, can determine which bit λ Bob sent. Hence the aim of the protocol is that the *overlaps*, the diagonal elements $G^{01}(A)$ and $G^{10}(A)$, relating the two different types of run, should decrease from 1 to 0, while the diagonal *weights* $G^{00}(A)$ and $G^{11}(A)$ remain equal to 1.

At this point it is worth noting that if both weights are not maintained—for example, if at the end $G^{00}(A) = 1$ while $G^{11}(A) = G^{01}(A) = 0$—Alice can still extract the value of λ by measuring whether or not the photon is in the state $|\psi^0\rangle$. Let us call this, for want of a better term, a *partial* protocol, in contrast to a *full* protocol that results in (9). A partial protocol can be used for one-way transmission, and the obvious advantage is that it costs nothing to transmit $\lambda = 1$. A possible disadvantage is that when Alice's measurement reveals nothing it could be because of some failure in the channel or in the measuring device. In the present discussion we focus on full protocols.

A very simple way to implement such a protocol is that on the very first step Alice sends the entire amplitude to Bob, with a Cost of 1 for this use of the channel. Bob then simply modifies this using the unitary $\mathcal{B}^\lambda$ and sends it back to Alice, either in one round or several rounds, with Alice sending nothing back. The Cost for using the channel in the Bob-to-Alice direction is also 1 (see the discussion in Sec. II B). Hence a total Cost of 2 for the protocol as a whole. Notice that since there is no restriction on $\mathcal{B}^\lambda$ this rather trivial protocol can be used to send quantum information. From the perspective of Cost, two-way protocols of the kind under discussion, in which initially all of the amplitude is on Alice's side, are interesting because a classical bit, $\lambda = 0$ or 1, can be sent from Bob to Alice at a total Cost of 1 rather than 2. And as shown below in Sec. III D, the product of the Costs for $\lambda = 0$ and 1 cannot be less than 1.

To discuss the successive steps in protocols that optimize the Cost, we need an appropriate notation. We will represent kets, thought of as column vectors, in the way suggested by the following example

$$|\psi\rangle = |a; c; b\rangle = |a_1, a_2, a_3; c_1, c_2; b_1, b_2\rangle, \qquad (10)$$

where the dimensions of the A, B, and C subspaces are $d(A) = 3$, $d(B) = 2$, and $d(C) = 2$, so $|\psi\rangle$ is an element of a seven-dimensional Hilbert space. Thus a_1, a_2, and a_3 are complex numbers forming a three-component vector a; similarly c and b are two-component vectors. If

$$|\hat{\psi}\rangle = |\hat{a}; \hat{c}; \hat{b}\rangle = |\hat{a}_1, \hat{a}_2, \hat{a}_3; \hat{c}_1, \hat{c}_2; \hat{b}_1, \hat{b}_2\rangle \qquad (11)$$

is another vector in the same space, its inner product with $|\psi\rangle$ is given by

$$\langle\hat{\psi}|\psi\rangle = \sum_{j=1}^{3} \hat{a}_j^* a_j + \sum_{k=1}^{2} \hat{c}_k^* c_k + \sum_{l=1}^{2} \hat{b}_l^* b_l. \qquad (12)$$

Note that we are dealing with a direct *sum* of subspaces $A \oplus B \oplus C$, *not* a tensor product of subsystems. In much of

what follows, B is empty or can be ignored, so $|a; c\rangle$ will suffice; this and other minor variants in notation should be self-explanatory.

Let us start with an extremely simple one-round full protocol with $d(A) = 2$ and $d(C) = 1$. It consists of the following steps:

$$|a_1, a_2; c\rangle = |1, 0; 0\rangle \to |1/\sqrt{2}, 0; 1/\sqrt{2}\rangle$$
$$\Rightarrow |1/\sqrt{2}, 0; (-1)^\lambda/\sqrt{2}\rangle \to |1/\sqrt{2}, (-1)^\lambda/\sqrt{2}; 0\rangle, \quad (13)$$

where 0 means this amplitude is equal to zero; do not confuse it with the label 0 for one of the two orthogonal states of a qubit. Here $\to$ indicates the action of a unitary on $A + C$ carried out by Alice and $\Rightarrow$ a λ-dependent unitary on C carried out by Bob. The action by Bob could involve intermediate steps requiring the B subspace, but its net effect is only to change the contents of C, so there is no need to include B in the discussion.

In words: At the outset all of the amplitude is in Alice's A, $a_1 = 1$. She maps half (in the sense of the absolute square) of it into C and sends it to Bob, who either sends it back unchanged in order $\lambda = 0$, or with the opposite phase to send $\lambda = 1$. Alice then empties the channel into the a_2 position, using a unitary on $A + C$ that is independent of λ, as it simply requires interchanging two subspaces. A final measurement by Alice determines which of the two orthogonal states is present in A and thus which bit Bob was sending.

Next consider what is happening to the Gram matrices $G^{\mu\nu}(A)$ and $G^{\mu\nu}(C)$ during the successive steps. In particular, the overlap $G^{01}(A)$ is equal to 1 at the outset, and the first step reduces it to $1/2$ by placing $+1/2$ in C. Bob's action changes $G^{01}(C)$ from $1/2$ to $-1/2$, and this negative contribution to the overlap moves back into A when Alice empties the channel, leading to the desired $G^{01}(A) = 0$. On the other hand, whereas the weight $G^{00}(A)$ is reduced to $1/2$ during the first step, Bob's action does not change the sign of $G^{00}(C)$, so in the final step Alice moves this weight back to its initial value of 1, and similarly for $G^{11}(A)$. Thus the goals of a full protocol have been achieved.

The Costs of using the channel are easily evaluated: $1/2$ for the Alice-to-Bob step and the same for the Bob-to-Alice step, for a total Cost of $Q^\lambda = 1$, the same for $\lambda = 0$ and 1. These satisfy the rigorous lower bound worked out below in Sec. III D, so this protocol is optimal if one uses total Cost as an appropriate measure of channel usage.

This protocol is easily extended to an equally efficient version involving N rounds, N any positive integer. Let

$$\epsilon = 1/2N \quad (14)$$

and for the first, $n = 1$, round replace (13) with

$$|1, 0; 0\rangle \to |\sqrt{1-\epsilon}, 0; \sqrt{\epsilon}\rangle$$
$$\Rightarrow |\sqrt{1-\epsilon}, 0; (-1)^\lambda \sqrt{\epsilon}\rangle \to |\sqrt{1-\epsilon}, (-1)^\lambda \sqrt{\epsilon}; 0\rangle,$$
$$(15)$$

while for round $n + 1$,

$$|\sqrt{1-n\epsilon}, (-1)^\lambda \sqrt{n\epsilon}; 0\rangle \to |\sqrt{1-(n+1)\epsilon}, (-1)^\lambda \sqrt{n\epsilon}; \sqrt{\epsilon}\rangle$$
$$\Rightarrow |\sqrt{1-(n+1)\epsilon}, (-1)^\lambda \sqrt{n\epsilon}; (-1)^\lambda \sqrt{\epsilon}\rangle$$
$$\to |\sqrt{1-(n+1)\epsilon}, (-1)^\lambda \sqrt{(n+1)\epsilon}; 0\rangle, \quad (16)$$

where it is straightforward to show that there exists a λ-independent unitary for the last step. The final result at the end of round N is the same as in (13), the case in which $N = 1$, and again the total Cost is $Q^0 = Q^1 = 1$, independent of λ. One can also let ϵ depend on n; thus $\epsilon_n > 0$ for round n, subject to the condition

$$\sum_n \epsilon_n = 1/2, \quad (17)$$

and the Cost is again equal to 1.

There are other protocols with larger Costs which may have some practical advantage. Thus, rather than a scalar amplitude, Alice might use photon polarization, say, horizontal H, which Bob could return as H to send $\lambda = 0$ or rotate to vertical V to send $\lambda = 1$. In this case the Costs are $Q^0 = Q^1 = 2$, so twice that for an optimal one-way protocol. However, there is now no need to maintain a particular phase relation between what is in Alice's domain and what is available to Bob during each round. If polarization is easier to maintain than phase—one leaves that up to the experts—one could imagine the added Cost being worthwhile if Alice has a large apparatus capable of generating single photons, while Bob, off on a trip to spy on Eve, needs only something easily carried in a suitcase.

The protocol used in SLAZ, in which Bob returns the amplitude for $\lambda = 0$, but absorbs it or feeds it to a measuring apparatus for $\lambda = 1$, looks less promising. Because the $\lambda = 1$ weight only moves from Alice to Bob it is difficult to have $G^{11}(A) = 1$ at the end of the protocol. In fact, SLAZ, discussed in Sec. IV, employs a clever trick (Zeno effect) to get around this problem, albeit at the cost of a large number of rounds to keep the probability of failure small and a large channel usage Cost for one of the bits.

D. Lower bound on Costs

The additivity and conservation properties of the Gram matrix $G^{\mu\nu}$ introduced in Sec. III A will now be used to obtain lower bounds on the total Cost of two-way protocols of the sort exemplified by, but not limited to, the case of one classical bit discussed above in Sec. III C. Using the $|a; c\rangle$ notation of (10)—the b entry is not needed in the following discussion—round n of an N-round protocol consists of the following steps carried out on $A + C$:

$$|a^\mu; 0\rangle_n \to |\bar{a}^\mu; c^\mu\rangle_n \Rightarrow |\bar{a}^\mu; \hat{c}^\mu\rangle_n \to |a^\mu; 0\rangle_{n+1}. \quad (18)$$

Here μ labels the bit which Bob is transmitting during this run. Thus, after Alice uses a unitary $\mathcal{A}_{n1}$ on $A + C$ to move some amplitude $|c^\mu\rangle_n$ into an initially empty channel, Bob applies a unitary $\mathcal{B}_n^\mu$ to $C + B$, leading to an amplitude $|\hat{c}^\mu\rangle_n$—note the circumflex (hat) added to c—in the channel. If Bob's action is passive, as assumed in Sec. III C (and in the later discussion of SLAZ in Sec. IV), one would have

$$\|\hat{c}^\mu\|_n \leqslant \|c^\mu\|_n, \quad (19)$$

but this conditions is actually not needed to obtain the general results and inequalities given below, which thus apply equally to one-way multitime transmission. As a final step Alice employs a unitary $\mathcal{A}_{n2}$ on $A + C$ to empty the channel by placing its amplitude into A. It is important that Alice's unitaries $\mathcal{A}_{n1}$

and $\mathcal{A}_{n2}$, unlike Bob's $\mathcal{B}_n^\mu$, *do not depend upon* μ, which can be different in different runs of the experiment.

The change in the Gram matrix associated with A during round n is given by

$$G_{n+1}^{\mu\nu}(A) - G_n^{\mu\nu}(A) = \langle a^\mu | a^\nu \rangle_{n+1} - \langle a^\mu | a^\nu \rangle_n$$
$$= \langle \hat{c}^\mu | \hat{c}^\nu \rangle_n - \langle c^\mu | c^\nu \rangle_n, \quad (20)$$

where $\langle a^\mu | a^\nu \rangle_n$ is the inner product of $|a^\mu\rangle_n$ and $|a^\nu\rangle_n$. The equality follows from the fact that $G^{\mu\nu}(A + C)$ is invariant under $\mathcal{A}_{n1}$ and $\mathcal{A}_{n2}$ and additive: $G^{\mu\nu}(A + C) = G^{\mu\nu}(A) + G^{\mu\nu}(C)$. To discuss the total change during N rounds, $n = 1, 2, \ldots N$, it is convenient to define

$$|C^\mu\rangle := \{|c^\mu\rangle_1, |c^\mu\rangle_2, \ldots, |c^\mu\rangle_N\},$$
$$|\hat{C}^\mu\rangle := \{|\hat{c}^\mu\rangle_1, |\hat{c}^\mu\rangle_2, \ldots, |\hat{c}^\mu\rangle_N\} \quad (21)$$

with inner products

$$\langle C^\mu | C^\nu \rangle = \sum_{n=1}^N \langle c^\mu | c^\nu \rangle_n, \quad \langle \hat{C}^\mu | \hat{C}^\nu \rangle = \sum_{n=1}^N \langle \hat{c}^\mu | \hat{c}^\nu \rangle_n. \quad (22)$$

Summing (20) over N rounds yields the formula

$$\Delta G^{\mu\nu}(A) = G_N^{\mu\nu}(A) - G_0^{\mu\nu}(A) = \langle \hat{C}^\mu | \hat{C}^\nu \rangle - \langle C^\mu | C^\nu \rangle \quad (23)$$

for the total change in the A portion of the Gram matrix during the full protocol. This quantity is bounded by

$$|\Delta G^{\mu\nu}(A)| \leqslant |\langle \hat{C}^\mu | \hat{C}^\nu \rangle| + |\langle C^\mu | C^\nu \rangle|$$
$$\leqslant \|\hat{C}^\mu\| \cdot \|\hat{C}^\nu\| + \|C^\mu\| \cdot \|C^\nu\| \quad (24)$$

using the norm $\langle C^\mu | C^\mu \rangle = \|C^\mu\|^2$.

Next define the total Cost K^μ for Alice-to-Bob and $\hat{K}^\mu$ for Bob-to-Alice uses of the channel, with Q^μ their sum:

$$K^\mu = \langle C^\mu | C^\mu \rangle = \|C^\mu\|^2, \quad \hat{K}^\mu = \langle C^\mu | C^\mu \rangle = \|\hat{C}^\mu\|^2,$$
$$Q^\mu = K^\mu + \hat{K}^\mu. \quad (25)$$

Combining (24) and (25) gives

$$|\Delta G^{\mu\nu}(A)| \leqslant \sqrt{K^\mu K^\nu} + \sqrt{\hat{K}^\mu \hat{K}^\nu} \leqslant \sqrt{Q^\mu Q^\nu}. \quad (26)$$

This yields an upper bound

$$\Delta G^{\mu\mu}(A) \leqslant Q^\mu \quad (27)$$

for a non-negative diagonal weight, and for the off-diagonal overlap

$$|\Delta G^{\mu\nu}(A)| \leqslant \sqrt{Q^\mu Q^\nu}. \quad (28)$$

In the particular case of the one-bit two-way protocol (Sec. III C), the aim is to reduce $G^{01}(A)$ from its initial value of 1 to 0 after N rounds. Setting $\mu = 0$ and $\nu = 1$ in (28), we see that to achieve this result it is necessarily the case that the Costs Q^0 and Q^1 for sending bits $\lambda = 0$ and $\lambda = 1$ must satisfy the condition

$$Q^0 Q^1 \geqslant 1. \quad (29)$$

This is satisfied as an equality with $Q^0 = Q^1 = 1$ for the specific protocols discussed in Sec. III C, which shows that they are optimal if total Cost is used as a measure. For more general protocols there is no reason to expect that the two Costs will be equal, and in that case if, say, the Cost for $\lambda = 1$ is made

very small, that for $\lambda = 0$ must be very large. This is in fact the case for the original SLAZ protocol, as discussed below in Sec. IV, which thus provides an interesting illustration of such a tradeoff.

IV. THE SLAZ PROTOCOL

A. Description of the protocol

The original SLAZ protocol differs from the simpler situation discussed in Sec. III C in two respects. First, it has a hierarchical structure: There are a large number M of outer rounds or cycles, each of which consists of a large number N of inner rounds or cycles, and the protocol will succeed with high probability provided

$$1 \ll M \ll N. \quad (30)$$

Second, while Bob sends a bit $\lambda = 0$ by reflecting the amplitude sent by Alice back into the channel, for $\lambda = 1$ he simply empties the channel, which can be described as a unitary operation in which the C amplitude is placed in Bob's subspace B. In addition, the original SLAZ protocol and some of its modifications involve measurements at intermediate times, and these will be replaced in the discussion below by unitary operations in the manner suggested at the end of Sec. II B.

We use a notation

$$|\psi\rangle = |a_1, a_2, a_3, a_4; c; b\rangle \quad (31)$$

of the form introduced in (10), where the a_j are scalar amplitudes in Alice's domain $A = A_1 + A_2 + A_3 + A_4$, c is the amplitude in the channel C, and b is in Bob's domain B. Here capital letters are used to denote subspaces and the corresponding projectors, while lowercase letters indicate (in general complex) scalar amplitudes. While A_4 and B are one dimensional, one can also make these larger spaces for reasons that will appear during the discussion. An abbreviated notation is often convenient: $|a_2, a_3\rangle$ in the case of a unitary acting on $A_2 + A_3$ while all the other amplitudes remain unchanged.

Central to the discussion are unitary operators that represent a rotation by an angle θ on a two-dimensional space:

$$R(\theta)|\alpha, \beta\rangle = |\alpha \cos\theta - \beta \sin\theta, \alpha \sin\theta + \beta \cos\theta\rangle. \quad (32)$$

In particular, R_M and R_N, defined in terms of small angles, play a central role:

$$R_M := R(\theta_M), \quad \theta_M := \pi/(2M); \quad R_N := R(\theta_N),$$
$$\theta_N := \pi/(2N). \quad (33)$$

Note in particular that

$$(R_M)^M = (R_N)^N = R(\pi/2); \quad R(\pi/2)|\alpha, \beta\rangle = |-\beta, \alpha\rangle. \quad (34)$$

In view of the fact that θ_N is a small angle, the following approximations turn out to be useful

$$\cos\theta_N \approx \exp\left[-\theta_N^2/2\right] = \exp[-\pi^2/(8N^2)] \approx 1 - \pi^2/(8N^2),$$
$$(\cos\theta_N)^N \approx \exp[-\pi^2/(8N)] \approx 1 - \pi^2/(8N) \approx 1, \quad (35)$$

and similarly if N is replaced by M.

These approximations are useful for understanding the overall structure of the protocol, which is the following. At

the beginning of outer round m, $1 \leqslant m \leqslant M$, R_M is applied to $A_1 + A_2$ to yield

$$|a_1, a_2\rangle^\lambda = R_M |\bar{a}_1, \bar{a}_2\rangle^\lambda, \qquad (36)$$

where $\bar{a}_1$ and $\bar{a}_2$ are the values of these amplitudes at the end of the previous outer round. In general they depend upon which bit $\lambda = 0$ or 1 is being transmitted, whence the superscript label, even though Alice's operations do not depend upon λ. The very first outer round $m = 1$ begins by applying (36) to the starting state (31) with $a_1 = 1$ and all the other amplitudes equal to zero.

The initial step (36) of outer round m is followed by a sequence of N inner rounds, each involving the following steps, here displayed using the type of notation employed in Sec. III C, but now with reference to the subspace $A_2 + A_3 + C$,

$$|a_2, a_3; c = 0\rangle \to |a'_2, a'_3; c = 0\rangle \to |a'_2, 0; a'_3\rangle$$
$$\Rightarrow |a'_2, 0; (1 - \lambda)a'_3\rangle \to |a'_2, (1 - \lambda)a'_3; 0\rangle, \qquad (37)$$

where

$$|a'_2, a'_3\rangle = R_N |a_2, a_3\rangle. \qquad (38)$$

In words, Alice applies the unitary rotation R_N (33) to $A_2 + A_3$ and then maps A_3 into the empty channel. Next comes Bob's action, indicated by $\Rightarrow$, to either reflect the amplitude a'_3 back into C if he is sending $\lambda = 0$ or shift it into his domain B, leaving the channel empty if sending $\lambda = 1$. Alice, who does not know the value of λ, maps whatever is in the channel back into A_3 by a unitary that simply exchanges the contents of A_3 and C and then begins the next inner round. The result of N inner rounds in succession is

$$|a_2, a_3\rangle \to \begin{cases} |0, a_2\rangle & \text{for } \lambda = 0 \\ |(\cos \theta_N)^N a_2, 0\rangle \approx |a_2, 0\rangle & \text{for } \lambda = 1, \end{cases} \qquad (39)$$

where the $\lambda = 1$ approximation is justified when N is very large [see (35)].

Following the N inner rounds Alice completes this outer round by applying a unitary to $A_3 + A_4$ that empties the contents of A_3 into A_4. For $\lambda = 1$, $a_3 = 0$ (39), so this emptying step is trivial, while for $\lambda = 0$ it is nontrivial and plays a significant role in understanding the true Costs of the protocol. In the original SLAZ protocol this emptying step is replaced by a measurement, but instead of a measurement one can just as well let the amplitudes accumulate in A_4, which is the perspective used here. At the end of the protocol after completing M outer rounds the final result is

$$\lambda = 0 : |a_1 = 1 - r_1, a_2 = 0, a_3 = 0, a_4 = r_4, c = 0, b = 0\rangle,$$
$$\lambda = 1 : |a_1 = s_1, a_2 = 1 - s_2, a_3 = 0, a_4 = 0, c = 0, b = s_b\rangle, \qquad (40)$$

where the quantities denoted by r_j and s_k are small corrections, of order $1/M$ or M/N. If these are ignored, all the amplitude is in A_1 for $\lambda = 0$ or A_2 for $\lambda = 1$, and a simple measurement allows Alice to determine which bit Bob sent.

B. Calculation of Costs and overlap

It is fairly straightforward to work out the Costs for the SLAZ protocol using approximations justified by $1 \ll M \ll N$, and the results are summarized in Sec. IV C below. We begin with the case $\lambda = 1$. If one ignores small quantities, the nonzero components of $|\psi\rangle_m$ at the beginning and at the end of outer round m are

$$a_1 = \cos(m\theta_M), \quad a_2 = \sin(m\theta_M), \qquad (41)$$

and since $M\theta_M = \pi/2$, at the end of outer round M the result is the $\lambda = 1$ line in (40).

The probability that the photon arrives in B during outer round m—the probability that Bob will detect it if he uses a measuring device—is the sum of the absolute squares of the amplitudes in the channel C in the N inner rounds, as this is an incoherent process:

$$N[\sin(m\theta_M)]^2[\sin(\theta_N)]^2 \approx (\pi^2/4)[\sin(m\theta_M)]^2/N. \qquad (42)$$

Summing over m gives the total probability

$$K^1 = Q^1 = (\pi^2/8)(M/N) \qquad (43)$$

that the photon will end up in Bob's domain by the end of the protocol, which is the same as the total Cost for $\lambda = 1$.

In the case $\lambda = 0$, any amplitude placed by Alice in C is immediately returned by Bob and at the end of each outer round is emptied into a_4, so at the end of outer round m the state is

$$|\psi\rangle_m = |a_1 = (\cos \theta_M)^m, a_2 = 0, a_3 = 0, a_4, c = 0, b = 0\rangle. \qquad (44)$$

For $m = M$ this is (40) with $r_1 = \pi^2/(8M)$. Thus at the end of the protocol a_2, a_3, c, and b are strictly zero. The Cost associated with inner round n—note that the channel is used twice—is

$$2[(\sin \theta_M) \sin(n\pi/2N)]^2 \qquad (45)$$

Summing over n gives a total of $(\pi^2/4)(N/M^2)$ for each outer round and hence for M outer rounds a total Cost of

$$Q^0 = (\pi^2/4)(N/M). \qquad (46)$$

To compute the total change in overlap $\Delta G^{01}(A)$, note that since for $\lambda = 1$ Bob does not return an amplitude; only the $\langle C^\mu | C^\nu \rangle$ term in (23) contributes. The contribution for inner round n of outer round m is the product of the factors

$$[\sin \theta_M \sin(n\theta_N)] \times [\sin(m\theta_M) \sin \theta_N] \qquad (47)$$

corresponding to $\lambda = 0$ and 1. Summing them yields

$$(\sin \theta_M \sin \theta_N) \sum_{m,n}^{M,N} \sin(m\theta_M) \sin(n\theta_N)$$
$$= (\pi^2/4MN)(4MN/\pi^2) = 1, \qquad (48)$$

and hence

$$\Delta G^{01}(A) = -1, \tag{49}$$

as expected.

C. Discussion of Costs and probabilities

Let us summarize the results of Sec. IV B. The total Costs Q^0 and Q^1 for $\lambda = 0$ and 1 are

$$Q^0 = (\pi^2/4)(N/M), \quad Q^1 = (\pi^2/8)(M/N),$$
$$Q^0 Q^1 \approx 3.044, \quad Q^0/Q^1 = 2N^2/M^2. \tag{50}$$

Given that $M \ll N$, Q^1 is minuscule, Q^0 is enormous, with their product is of order 1 and satisfies the rigorous bound (29). The case $\lambda = 1$ is the easiest to understand. Since Bob does not return the amplitude put into the channel by Alice, the Bob-to-Alice Cost $\hat{K}^1$ is zero. The Alice-to-Bob Cost is $|s_b|^2$ in (40), i.e., the probability that at the very end the photon is in Bob's domain. The physical reason for this is that the process by which the amplitude gets there is *incoherent* (no quantum interference) since no amplitude goes back through the channel. Bob could either accumulate these amplitudes until the end of the protocol and then measure to see if the photon is in B or carry out a measurement at the end of each inner round; in either case the probability of his detecting the photon is $|s_b|^2$ in (40). The situation is analogous to the use of intermediate time measurements in a one-way protocol as discussed at the end of Sec. II B.

The enormous Cost Q^0 for $\lambda = 0$ comes about because Bob repeatedly returns the amplitude sent by Alice in a coherent process. While the amplitude bouncing back and forth through the channel is relatively small, of order $1/M$, multiplying its absolute square by $2N$, the number of times this amplitude is in the channel during each outer round, leads to a Cost of order N/M^2 for each outer round and hence a total of order N/M for the complete process.

Clearly, the large value of Q^0 means the claim that the protocol is counterfactual cannot be maintained if Cost is used as a criterion for channel use, so it is worth discussing how the authors of SLAZ reached a different conclusion. In essence, their reasoning was based on the small value of the amplitude in A_3 at the end of an outer round just before it is transferred to A_4, as per the discussion in Sec. IV A. The absolute square of this amplitude is the probability that the corresponding detector D_3 in Fig. 2(b) in the SLAZ paper will be triggered. This amplitude was earlier oscillating back and forth inside the subspace with projector $S = A_2 + A_3 + C$, and hence it is reasonable to assume that if this detector triggers, the photon was earlier in S during all N inner rounds making up this particular outer round.[1] As this probability is of order $1/M^2$, the probability that one of the D_3 detectors triggers during the M outer rounds that make up a given run is of order $1/M$ and hence small.

There are two serious objections to using this small probability to justify the claim that the protocol is counterfactual:

_____

[1]This assumption can be justified using the Consistent Histories approach; see the discussion of measurements in [5,6].

one classical and the other quantum. Let us start with the former. During a particular outer round the photon amplitude in a $\lambda = 0$ run rattles back and forth inside S a total of N times, and in particular it is in C a total of $2N$ times. Consider a stochastic classical protocol for transmitting information in which most of the time Alice and Bob exchange no information at all. However, with a small probability ϵ Alice sends a little white ball into the channel leading to Bob, who colors it green or red and sends it back to Alice to convey one bit of information. She records the color, paints the ball white, and returns it to Bob who again colors it to send a second bit, and so forth, for a total of N rounds. The average rate of transmitting information is $N\epsilon$ bits, and one cannot simply throw away the factor of N and claim that this protocol is in some sense counterfactual.

The quantum difficulty has to do with what can be inferred from the probability that the photon was in $S = A_2 + A_3 + C$ during the inner rounds that make up a particular outer round. One may be tempted to use classical reasoning and assume that the probabilities of being in each of the mutually exclusive regions A_2, A_3, and C that combine to make up S are well defined and sum to the probability of being in S. However, in the presence of quantum interference this sort of reasoning is invalid and leads to paradoxes. (See the discussion of parallel channels in Sec. II A.)

V. CONCLUSION

The original SLAZ proposal has motivated a large number of papers (see the extensive bibliographies in [3,4]). Merely trying to summarize them, much less provide a detailed review, lies outside the scope of the present paper. Broadly speaking, this literature consists of modifications, extensions, or improvements of the original SLAZ scheme, along with criticisms of the claim that these protocols are counterfactual and replies to such criticisms. It is hoped that the following rather brief comments will provide some orientation.

Significant extensions of the original SLAZ scheme by the last three members of the original collaboration include the use of a phase change rather than absorption to transmit the $\lambda = 1$ bit [7], a scheme to transmit quantum states by multiple iterations of the original SLAZ scheme [8], and using many photons in place of a single photon to transmit a classical bit [4]. These and others are certainly interesting ideas from the perspective of transmitting quantum information and worth further exploration.

On the other hand, in these and all other extensions or modifications of SLAZ this author has examined, the claim that the protocol is counterfactual, in the sense that the total use of a quantum channel is negligible in the asymptotic limit, is subject to the same objections discussed in Sec. IV C: An incorrect use of probabilistic reasoning in a situation where quantum interference means probabilities cannot be defined and where even in a classical situation Cost would be better than probability as a measure of channel usage. The total Cost remains finite in the asymptotic limit of a very large number of steps, which means that counterfactual claims should be dropped. Doing so will aid, not hinder, the serious study of these interesting quantum schemes for transmitting information.

Shortly after the original SLAZ publication, Vaidman published a Comment [9] claiming that in the $\lambda = 0$ case in which Bob reflects the amplitude rather than absorbing it, the photon which was later (with high probability) detected by Alice must at an earlier time have been in the channel C. In their Reply [10] the SLAZ authors pointed out this way of reasoning about events at an intermediate time in the presence of quantum interference was invalid and leads to paradoxes, a position supported by the analysis in Sec. IV B above. However, they then repeated their original counterfactual claim which itself is based on a defective understanding of probabilities about probabilities at an intermediate time. A later and much more extended criticism of counterfactuality claims by Vaidman [11] suffers from the same difficulty as his earlier Comment.

Some years later Aharonov and Vaidman [12] claimed to have found a scheme of the general SLAZ type which is genuinely counterfactual. However, when measurements or absorption of a photon at intermediate times are replaced by unitary processes—mapping amplitude into an empty subspace reserved for this purpose, as discussed in Sec. IV A—the inequality in Sec. III D applies to this case and undermines the counterfactual claim. The fundamental difficulty with such claims is that the Hilbert-space projector which identifies the position of a particle at some intermediate time does not commute with the one representing the quantum state evolving unitarily in time.

The most significant contributions of the present paper to the analysis of SLAZ-type protocols is the use of Cost as a measure of channel usage and the use of Gram matrices for discussing information transfer at intermediate times in the presence of quantum interference. In particular, the fact that these Gram matrices are additive over subspaces and invariant (conserved) under unitary time transformations plays a key part in the discussion in Sec. III. A rather surprising feature is the role of off-diagonal elements, overlaps, as a type of information measure which, unlike most such measures, is not in general positive. That it can be negative plays a very significant part in understanding its intuitive role in information transfer. That its total change on Alice's side must be -1 during the course of a successful protocol is confirmed for the SLAZ protocol in Sec. IV B.

This use of Gram matrices requires that the intermediate time steps be unitary. In the case of SLAZ, measurements at intermediate times can be eliminated by mapping photon amplitude into empty subspaces, and this can be achieved in certain other cases, e.g., the Aharonov and Vaidman protocol [12]. However, it is less clear whether something similar could be done in a case in which, for example, Alice uses measurements at intermediate times to change later steps in the protocol in hopes of reducing the total Cost. This author believes that such an improvement is impossible, because measurements themselves are quantum processes whose description simply requires a large enough Hilbert space in Alice's domain [13]. However, this has not yet been demonstrated.

What is special about classical information? Sending an arbitrary one-qubit quantum state from Alice to Bob using the two-way protocol of Sec. III C could be done with a Cost of 2, which is to say twice that of simply using a one-way protocol from Bob to Alice. That this is the minimum seems likely, but has not been demonstrated. What about a two-way protocol with all the amplitude starting on Alice's side, with the aim of a perfect transmission of each of two specified *nonorthogonal* states from Bob to Alice – what would be the minimum total Cost?

An interesting feature of the original SLAZ protocol is the enormous ratio $2N^2/M^2$ [see (50)] of the Costs to transmit $\lambda = 0$ and 1, in contrast to the relatively simple protocols discussed in Sec. III C for which the ratio is 1. Because the success of SLAZ depends upon N being much larger than M, this large ratio presumably has something to do with Bob's not sending anything back through the channel when $\lambda = 1$. Might there be some interesting physical principles, in addition to the Zeno effect, hiding here and waiting to be explored?

In conclusion, it is hoped that the thinking and tools employed in this paper will be useful for studying other problems of quantum information at intermediate times in situations where the careless use of ill-defined probabilities generates paradoxes rather than physical understanding. In particular, information transfer among three or more parties, of current interest in the study of quantum networks, might benefit from the sort of analysis used here.

[1] H. Salih, Z.-H. Li, M. Al-Amri, and M. S. Zubairy, Protocol for Direct Counterfactual Quantum Communication, Phys. Rev. Lett. **110**, 170502 (2013).

[2] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin, 1932) [R. T. Beyer, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, 1955)].

[3] J. R. Hance, J. Ladyman, and J. Rarity, How quantum is quantum counterfactual communication? Found. Phys. **51**, 12 (2021).

[4] Z.-H. Li, S.-Y. Feng, M. Al-Amri, and M. Suhail Zubairy, Direct counterfactual quantum communication protocol beyond a single photon source, Phys. Rev. A **106**, 032610 (2022).

[5] R. B. Griffiths, What quantum measurements measure, Phys. Rev. A **96**, 032110 (2017).

[6] R. B. Griffiths, The Consistent Histories approach to quantum mechanics, in *Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta (Stanford University, Stanford, 2019), available at https://plato.stanford.edu/entries/qm-consistent-histories/.

[7] Z.-H. Li, M. Al-Amri, and M. S. Zubairy, Direct quantum communication with almost invisible photons, Phys. Rev. A **89**, 052334 (2014).

[8] Z.-H. Li, M. Al-Amri, and M. S. Zubairy, Direct counterfactual transmission of a quantum state, Phys. Rev. A **92**, 052315 (2015).

[9] L. Vaidman, Comment on "Protocol for Direct Counterfactual Quantum Communication", Phys. Rev. Lett. **112**, 208901 (2014).

[10] H. Salih, Z.-H. Li, M. Al-Amri, and M. S. Zubairy, Salih *et al.* Reply, Phys. Rev. Lett. **112**, 208902 (2014).

[11] L. Vaidman, Counterfactuality of 'counterfactual' communication, J. Phys. A: Math. Theor. **48**, 465303 (2015).

[12] Y. Aharonov and L. Vaidman, Modification of counterfactual communication protocols that eliminates weak particle traces, Phys. Rev. A **99**, 010103(R) (2019).

[13] For a consistent quantum-mechanical description of the measuring process, see [5], the relevant sections of [6], and Chaps. 17 and 18 of [14].

[14] R. B. Griffiths, *Consistent Quantum Theory* (Cambridge University Press, Cambridge, 2002).