Optimal quantum operations at zero energy cost

Giulio Chiribella[1,2,3] and Yuxiang Yang[1,2]

[1]*Department of Computer Science, University of Hong Kong, Pokfulam Road, Hong Kong, China*
[2]*University of Hong Kong Shenzhen Institute of Research and Innovation, Kejizhong 2nd Road, Shenzhen, China*
[3]*Canadian Institute for Advanced Research, CIFAR Program in Quantum Information Science, Toronto, Ontario M5G 1Z8, Canada*

Quantum technologies are developing powerful tools to generate and manipulate coherent superpositions of different energy levels. Envisaging a new generation of energy-efficient quantum devices, here we explore how coherence can be manipulated without exchanging energy with the surrounding environment. We start from the task of converting a coherent superposition of energy eigenstates into another. We identify the optimal energy-preserving operations, both in the deterministic and in the probabilistic scenario. We then design a recursive protocol, wherein a branching sequence of energy-preserving filters increases the probability of success while reaching maximum fidelity at each iteration. Building on the recursive protocol, we construct efficient approximations of the optimal fidelity-probability trade-off, by taking coherent superpositions of the different branches generated by probabilistic filtering. The benefits of this construction are illustrated in applications to quantum metrology, quantum cloning, coherent state amplification, and ancilla-driven computation. Finally, we extend our results to transitions where the input state is generally mixed and we apply our findings to the task of purifying quantum coherence.

I. INTRODUCTION

Rapid experimental advances are pushing towards the realization of new quantum technologies [1–6]. Decoherence still remains the grand challenge, but, as quantum technologies approach real-life applications, questions of energy efficiency are bound to become increasingly more relevant. Nowadays, energy efficiency is one of the major problems in information and communication technology [7] and, as such, it is the object of a large amount of research, both experimentally [8,9] and theoretically [10,11]. In this area, quantum technologies hold a large, relatively unexplored potential, which is likely to become critical in the long-term future. In this perspective, it is compelling to explore the ultimate performances achieved by quantum devices with limited energy resources. The problem is not only of fundamental interest. Pioneering experiments in quantum optomechanics have already started to develop the tools for manipulating quantum systems with minimal amounts of energy [12]. Similarly, engineered light-matter interactions in quantum dots [13] and superconducting circuits [14] enable the control of dynamics at the level of single quanta, offering a promising platform for the realization of prototypes of energy-optimized quantum devices.

In order to address the question of energy efficiency, one needs to characterize the quantum operations that can be performed with given energy resources. Concretely, an energy resource is described by the state of a battery, i.e., an auxiliary system that exchanges energy with the system used as the information register. The constraint that the battery is the *only* energy resource used in the processing amounts to the requirement that the joint evolution of register and battery be *energy-preserving*. In general, energy-preserving evolutions need not be reversible: the register and the battery can interact nontrivially with auxiliary degrees of freedom, as long as they do not exchange energy with it. This scenario is illustrated in Fig. 1(a). The achievable operations are then modeled as reduced evolutions of the information register, with the battery

initialized in a given resource state. In this model, characterizing and optimizing the energy-preserving operations on the composite system of battery and information register is an essential step towards characterizing and optimizing the achievable operations on the information register alone.

The focus of this paper is the characterization and optimization of energy-preserving operations, which provide the foundation to the broader program aimed at designing energy-optimized quantum devices. Energy-preserving operations are also interesting *per se*, as the operations that can be realized without the assistance of external energy resources. This property is appealing in situations where a device has to switch to a "low-power mode," as it is often the case for nowadays mobile devices and will arguably be the case also for future devices incorporating quantum gadgets [15]. In these situations, it may become convenient to disconnect part of the device from the battery and to let that part of the device operate in an energy-autonomous way, as illustrated in Fig. 1(b). Studying energy-preserving operations is also a useful strategy to derive results about other related types of operations. For example, the energy-preserving family includes as a special case the operations that can be achieved with passive optical elements like beam splitters, phase shifters, and mirrors. As a result, the optimization of a desired task over all energy-preserving operations yields upper bounds on the performance of arbitrary quantum circuits built with passive optical elements. Results about energy-preserving operations can also be used in situations involving the preservation of observables other than the energy. This is an important point, because constraints on the preservation of suitable observables occur naturally in many applications. An interesting example arises in quantum dots, where the implementation of logical gates often benefits from the existence of "sweet spots"— special working points where the charge noise is suppressed. The set of gates that can be performed at the sweet spot is limited: for example, in the three-electron exchange-only spin qubit [16], only the rotations around the z axis can be
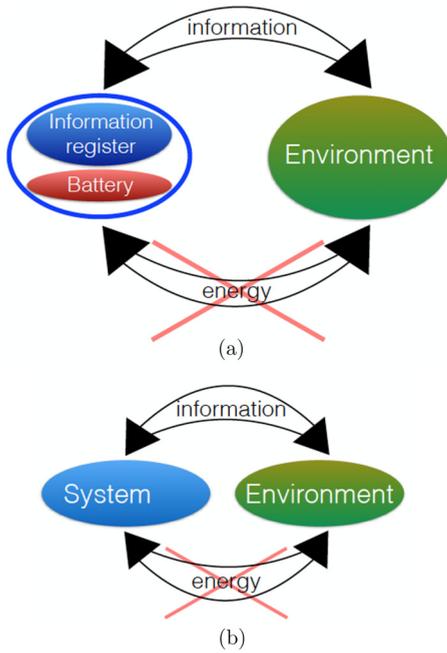
FIG. 1. Manipulating quantum information with limited energy resources. (a) Quantum device powered by a battery. The device contains an information register, where data are stored, and a battery, representing the energy resource(s) used to operate on the data. The information register and the battery are allowed to interact with the surrounding environment (possibly including ancillas), as long as the interaction involves no exchange of energy. The situation where the device uses energy from the environment to aliment its battery can be included in the model by formally regarding the energy sources in the environment as part of the battery. (b) Quantum device in low-power mode. In this mode, some quantum operations are performed without the aid of the battery, i.e., relying only on the interaction between the information register (here denoted simply as "the system") and the environment (possibly including ancillas). The evolution of the system is described by an energy-preserving operation.

performed at the sweet spot, while all other rotations incur into undesired noise [17]. Regarding the component of the spin along the z axis as the "energy," it follows that the "energy-preserving" channels are exactly the operations that can be performed with suppressed noise. Similar physics arises in superconducting flux qubits, where the sweet spot is with respect to magnetic noise [18].

This paper characterizes the set of energy-preserving operations and identifies the most efficient strategies for the manipulation of quantum states exhibiting coherence across energy levels. We will start from the basic task of transforming a pure superposition of energy eigenstates into another pure superposition. In this context, it is interesting to consider not only deterministic operations but also probabilistic operations arising from measurements that can be implemented at zero energy cost. By allowing for a nonunit probability of success, we find out that the constraint of energy preservation can be stretched to a previously unsuspected extent. For example, we will see that a beam of N atoms, each of them prepared in the superposition $|S\rangle = (|E\rangle + |G\rangle)/\sqrt{2}$ of the ground state and

the first excited state, can be probabilistically transformed at no energy cost into a stronger beam of $N^{2-\epsilon}$ atoms in a state that is nearly identical to the state of $N^{2-\epsilon}$ identical copies of $|S\rangle$, up to an exponentially small error. The ability to efficiently approximate forbidden transformations of coherence at zero energy cost is a new twist of the postselection approach widely applied in quantum information [19–32] and complements existing results on the resource theory of coherence [33–39].

After having characterized the structure of the optimal energy-preserving operations, we move to a different scenario, where the probability of success is not fixed *a priori*. We consider adaptive protocols whereby the experimenter performs repeated rounds of probabilistic operations and is free to decide on the fly whether to be content with the result obtained so far or whether to continue further. In this scenario we design a recursive protocol, consisting of a sequence of energy-preserving binary measurements that produce at each step the best approximation of the target with the highest probability allowed by quantum mechanics. Subsequent iterations of our protocol lead to an increasing probability of success, but also to a degradation of the state of the system and, eventually, to the loss of the advantages of postselection. This behavior is a consequence of the inevitable trade-off between the enhanced performance of probabilistic transformations and their reduced probability of success. The advantage of our recursive protocol is that it gives an explicit, ready-to-apply method to construct lower bounds to the optimal trade-off curve between fidelity and probability of success, a problem that so far has been solved only in one case [26]. Remarkably, in this particular case we find out that our protocol reproduces the optimal trade-off curve, provided that the input state is a superposition of sufficiently many energy levels. We conjecture that our protocol is asymptotically optimal also in those situations where the input and the target consist of many identical copies of pure states, an example being the asymptotic cloning of quantum coherence. Independently of the validity of this conjecture, the importance of the recursive protocol can be best appreciated in all those cases where the optimal trade-off curve is not explicitly known. To get even better lower bounds to the optimal curve, we finally introduce the operation of *coherent coarse-graining*, which consists in joining different outcomes into a single quantum operation. Coherent coarse-graining allows one to keep the same probability of success of the outcomes that are joined, while increasing the fidelity with the target. Combined with our recursive protocol, this operation provides a canonical way to generate analytical lower bounds to the optimal trade-off between fidelity and success probability, whose exact calculation is generally hard to perform without resorting to numerical optimization.

To demonstrate the broad applicability of our methods, we illustrate the recursive protocol and its coherent coarse-graining in a number of concrete examples, including quantum phase estimation, energy-preserving amplification of coherent states, and the optimal design of correction operations for ancilla-driven computation. In addition to the applications presented explicitly in the paper, our results can be directly applied to most of the canonical problems of quantum information processing, such as optimal state discrimination, gate programming, entanglement conversion, universal NOT,

and universal transpose, whose implementation is significantly affected once one imposes the requirement that no energy should be drawn from the environment.

Finally, we go beyond the regime of pure state transitions and extend our results to transitions where mixed states are given as inputs. For such transitions, we prove upper bounds for the performances of both deterministic and probabilistic operations, providing conditions for the saturation of the bounds. This extension provides optimal strategies for the implementation of quantum tasks such as purification [32,40,41] and broadcasting of mixed states [42].

The paper is organized as follows. Section II introduces the basic framework. In Sec. III we characterize the optimal energy-preserving process. Using this result, we construct the recursive protocol and study the operation of coherent coarse-graining in Sec. IV and apply it to several tasks in quantum information processing (Sec. V), including phase estimation (Secs. V A and V B), state cloning (Sec. V C), coherent light amplification (Sec. V D), and ancilla-driven computation (Sec. V E).The extension of our results to mixed states is discussed in Sec. VI. Finally, the conclusions are drawn in Sec. VII.

II. THE ENERGY-PRESERVING PARADIGM

In this section we introduce the framework that will be adopted in the rest of the paper. We first present the class of energy-preserving channels, which model deterministic evolutions that can be implemented without drawing energy from the environment. We then move to probabilistic operations, characterizing the stochastic evolutions that can be implemented at zero energy cost.

A. Energy-preserving channels

Consider a quantum system interacting with the surrounding environment from time t_1 to time t_2 through an interaction Hamiltonian $H_{int}(t)$, which we assume to be zero before t_1 and after t_2. As a result of the interaction, the system and the environment evolve jointly via the unitary operator

$$U = \mathcal{T} \exp \left\{ -\frac{i}{\hbar} \int_{t_1}^{t_2} dt [H_{sys} + H_{env} + H_{int}(t)] \right\}, \quad (1)$$

where $\mathcal{T}$ exp denotes the time-ordered exponential and H_{sys} and H_{env} are the Hamiltonians of the system and of the environment, both assumed to be time-independent. Regarding energy as a resource, we require the evolution to satisfy the condition

$$U^\dagger (H_{sys} + H_{env}) U = H_{sys} + H_{env}, \quad (2)$$

meaning that the total energy after t_2 is equal to the total energy before t_1. A sufficient condition for the validity of Eq. (2) is the commutation relation

$$[H_{int}(t), H_{sys} + H_{env}] = 0 \quad \forall t \in [t_1, t_2], \quad (3)$$

which guarantees that the sum of the system and environment energies is a constant of motion during the entire evolution. Our analysis covers this case and, possibly, more general scenarios where the sum of the system and environment energies is not a constant of motion at all times: condition (2) is generally

weaker than condition (3). Note however that conditions (2) and (3) define the same set of input-output evolutions from time t_1 to time t_2: for every unitary U that commutes with $H_{sys} + H_{env}$ one can always find a suitable interaction Hamiltonian $H'_{int}(t)$ that generates U as and commutes with $H_{sys} + H_{env}$ at all times.

Among the evolutions that conserve the total energy, we are interested in those that leave the energy of the environment untouched. Such evolutions satisfy the additional condition

$$U^\dagger H_{env} U = H_{env}. \quad (4)$$

Clearly, the combination of Eqs. (2) and (4) implies that the evolution has to preserve the energy of the system, namely

$$U^\dagger H_{sys} U = H_{sys}. \quad (5)$$

Assuming that the environment is initially in the state ρ_{env}, the effective evolution of the system is described by the quantum channel (completely positive trace-preserving map) $\mathcal{M}$ defined by

$$\mathcal{M}(\rho) = \mathrm{Tr}_{env}[U(\rho \otimes \rho_{env})U^\dagger] \quad (6)$$

where Tr_{env} denotes the partial trace over the Hilbert space of the environment. By construction, the channel $\mathcal{M}$ preserves the expectation value of the system's energy. Even more strongly, condition (5) implies that the channel $\mathcal{M}$ preserves also the variance and all the momenta of the Hamiltonian, namely

$$\mathcal{M}^\dagger (H_{sys}^n) = H_{sys}^n \quad \forall n \in \mathbb{N}, \quad (7)$$

where $\mathcal{M}^\dagger$ is the completely positive identity-preserving map describing the evolution in the Heisenberg picture, defined by

$$\mathcal{M}^\dagger(A) = \mathrm{Tr}_{env}[(I_{sys} \otimes \rho_{env})U^\dagger (A \otimes I_{env})U], \quad (8)$$

for arbitrary operators A. When Eq. (7) is satisfied, we say that $\mathcal{M}$ is an *energy-preserving channel*.

The energy-preserving condition (7) is equivalent to the requirement that the evolution $\mathcal{M}$ does not affect the probability distribution of the energy, i.e., that one has the equality of probabilities

$$\mathrm{Tr}[P_E \mathcal{M}(\rho)] = \mathrm{Tr}[P_E \rho], \quad (9)$$

where ρ is an arbitrary state, E is an arbitrary eigenvalue of H_{sys}, and P_E is the projector on the eigenspace corresponding to E.

It is easy to see that every energy-preserving channel $\mathcal{M}$ is *covariant* with respect to the free time evolution of the system; that is, $\mathcal{M}$ satisfies the condition

$$\mathcal{M}(U_t \cdot U_t^\dagger) = U_t \mathcal{M}(\cdot) U_t^\dagger \quad \forall t \in \mathbb{R}, \quad (10)$$

with $U_t = \exp[-it H_{sys}/\hbar]$. However, the converse is not true in general: for example, a channel that discards the input state and re-prepares an eigenstate of the energy is covariant but not energy-preserving. Physically, the difference between energy-preserving and covariant channels is that covariant channels preserve the sum $H_{sys} + H_{env}$, while the energy-preserving channels separately preserve the individual summands H_{sys} and H_{env}. Further discussion on the relation between energy-preserving channels and other classes of channels, such as incoherent channels and Hadamard channels, is presented in Appendix A.

Energy-preserving channels provide an economical way to process information in situations where energy becomes a scarce resource. For example, one can imagine a scenario where a microscopic machine has to perform a desired task on the system (like copying data, correcting for an error, or erasing information) without changing the energy of the surrounding environment. While scenarios of this kind are not a practical reality yet, prototype demonstrations of energy-preserving channels can be conceived for quantum systems with a high degree of control, such as ion traps [43,44], optical cavities [45,46], optical lattices [47,48], and optomechanical systems [49,50].

B. Structure of the energy-preserving channels

We have seen that every energy-preserving interaction induces an energy-preserving quantum channel, i.e., a channel satisfying Eq. (7). The converse is also true: given an energy-preserving channel $\mathcal{M}$, one can always engineer an interaction between the system and the environment that conserves the total energy and does not draw energy from the environment. To establish this fact, note that the map $\mathcal{M}^\dagger$ satisfies the condition

$$\mathcal{M}^\dagger(A) = A \quad \forall A \in \mathcal{A}, \tag{11}$$

where $\mathcal{A}$ is the commutative algebra generated by the powers of the Hamiltonian. The algebra $\mathcal{A}$ contains the identity and is closed under adjoint. Technically, algebras of this kind are known as unital $*$-algebras [51]. For any such algebra, the maps that satisfy Eq. (11) are characterized by a simple lemma:

Lemma 1 (Lindblad [52]). Let $\mathcal{M}^\dagger$ be an identity-preserving completely positive map, written in the Kraus form $\mathcal{M}^\dagger(A) = \sum_{k=1}^{K} M_k^\dagger A M_k$, and let $\mathcal{A}$ be a unital $*$-algebra $\mathcal{A}$. The map $\mathcal{M}^\dagger$ preserves the elements of $\mathcal{A}$ if and only if each Kraus operator M_k belongs to the commutant of $\mathcal{A}$, i.e., to the set of operators

$$\mathcal{A}' := \{B \in \mathcal{B}(\mathcal{H}), [A,B] = 0 \quad \forall A \in \mathcal{A}\},$$

$\mathcal{B}(\mathcal{H})$ denoting the set of bounded operators on $\mathcal{H}$.

In the case of the Abelian algebra generated by the powers of H_{sys}, the commutation condition reduces to

$$[M_k, H_{\text{sys}}] = 0, \tag{12}$$

meaning that each Kraus operator M_k must be of the block diagonal form

$$M_k = \bigoplus_E P_E M_k P_E, \tag{13}$$

with the sum running over the eigenvalues of H_{sys}. All throughout the paper we will assume the energy spectrum to be discrete.

As a consequence of the block diagonal form (13), one can realize the channel $\mathcal{M}$ through an energy-preserving isometry. Specifically, one can express the channel as

$$\mathcal{M}(\rho) = \text{Tr}_{\text{env}}[V \rho V^\dagger], \tag{14}$$

where V is the isometry defined by

$$V : \mathcal{H}_{\text{sys}} \to \mathcal{H}_{\text{sys}} \otimes \mathcal{H}_{\text{env}}, \quad V := \sum_{k=1}^{K} M_k \otimes |\phi_k\rangle, \tag{15}$$

$\{|\phi_k\rangle\}_{k=1}^{K}$ being a set of orthonormal states in the environment's Hilbert space. With this definition, the isometry V satisfies the relation

$$V H_{\text{sys}} = (H_{\text{sys}} \otimes I_{\text{env}})V. \tag{16}$$

In turn, Eq. (16) implies that the isometry V can be realized via an interaction that preserves both the energy of the system and the energy of the environment:

Theorem 1. Let $\mathcal{M}$ be a quantum channel transforming states on $\mathcal{H}$. Then, the following are equivalent:

(1) $\mathcal{M}$ is energy-preserving;

(2) $\mathcal{M}$ can be realized through a joint evolution of the system together with an environment of the following form:

$$\mathcal{M}(\rho) = \text{Tr}_{\text{env}}[U(\rho \otimes |\phi_0\rangle\langle\phi_0|)U^\dagger], \tag{17}$$

where $|\phi_0\rangle$ belongs to ground eigenspace of H_{env} and U is a unitary evolution that commutes with H_{sys} and with H_{env};

(3) $\mathcal{M}$ can be realized through a joint evolution as in Eq. (17), where the unitary U is generated by an interaction $H_{\text{int}}(t)$ that commutes with H_{sys} and H_{env} at all times.

The proof is presented in Appendix B.

Note the appearance of the Hamiltonian of the environment in the statement of the theorem. This is natural, because in general we are dealing with the evolution of an open system. In this situation, the Hamiltonian of the environment plays a crucial role, for it determines the minimum amount of energy one has to invest into the realization of the desired channel. Ideally, Theorem 1 guarantees that such amount can be reduced to zero in the case of energy-preserving channels. Specifically, the desired evolution can be engineered by initializing the environment in an eigenstate of its Hamiltonian and by turning on a coupling that preserves the individual energies of system and environment, keeping the latter inside the ground space for the whole time evolution. As a result, the evolution is implemented at zero energy cost, at the price of an entropy increase in the environment, which is generally left in a mixture of states with the same energy. In other words, the environment is only used passively as computational workspace wherein information can be stored.

C. Energy-preserving instruments

While in the case of deterministic evolutions the notion of energy preservation is pretty straightforward, the situation is different for probabilistic transformations induced by quantum measurements. In this section we introduce a notion of *probabilistic energy-preserving transformations*, which characterizes those operations that can be implemented (in principle) without paying an energy cost.

Adopting Ozawa's model of the measurement process [53], we view probabilistic evolutions as the result of a unitary interaction between the system and the environment, followed by the projective measurement of a "meter observable" O on the environment. In this model, the preservation of the energy imposes constraints on the interaction as well as constraints on the measurement. Like in the previous sections, we demand that the system-environment interaction preserve the total energy of the system and the environment [Eq. (2)] and do not change the energy of the environment [Eq. (4)]. As we argued in the previous paragraph, these requirements characterize

the evolutions that can be implemented at zero energy cost. Regarding the measurement, we demand that it should not disturb the energy of the environment, or, equivalently, that the meter observable and the energy are compatible observables, namely

$$[O, H_{\text{env}}] = 0. \tag{18}$$

If this condition were not satisfied, the measurement of O would collapse an eigenstate of the energy into a random eigenstate of O, thus altering the expectation value of the energy. Observables that commute with the Hamiltonian of the environment are the prototype of measurements that can be performed without paying an energy cost. The conditions (2), (4), and (18) are the standard requirements put forward in the Wigner-Araki-Yanase theorem [54–56] and in all the works that followed it up [57–60]. In this context the commutation relation (18) is known as *Yanase's condition* [56]. Recently, the same framework discussed here has been used as the starting point to define energy requirements for the implementation of quantum measurements [60,61].

Let us analyze the probabilistic evolutions resulting from the requirement of zero energy cost. According to quantum measurement theory [53,62–64], the measurement of O induces a stochastic evolution of the state of the system, described by a *quantum instrument*, namely a collection of quantum operations (completely positive trace nonincreasing maps) $\{\mathcal{M}_x\}_{x \in \mathsf{X}}$ subject to the normalization condition

$$\sum_{x \in \mathsf{X}} \text{Tr}[\mathcal{M}_x(\rho)] = \text{Tr}[\rho] \tag{19}$$

for every quantum state ρ. For a system prepared in the state ρ, the measurement generates the outcome $x \in \mathsf{X}$ with probability

$$p(x|\rho) = \text{Tr}[\mathcal{M}_x(\rho)], \tag{20}$$

and, conditionally on outcome x, returns the system in the state

$$\rho_x' = \frac{\mathcal{M}_x(\rho)}{\text{Tr}[\mathcal{M}_x(\rho)]}. \tag{21}$$

In the model considered here, the set of outcomes is the spectrum of the meter observable O and the quantum operation $\mathcal{M}_x$ is defined by

$$\mathcal{M}_x(\rho) = \text{Tr}_{\text{env}}[(I_{\text{sys}} \otimes Q_x)U(\rho \otimes \rho_{\text{env}})U^\dagger], \tag{22}$$

Q_x being the projector on the eigenspace of O with eigenvalue x. Note that, by summing over all possible outcomes, one obtains

$$\sum_{x \in \mathsf{X}} \mathcal{M}_x(\rho) = \text{Tr}_{\text{env}}[U(\rho \otimes \rho_{\text{env}})U^\dagger] \equiv \mathcal{M}(\rho),$$

where $\mathcal{M}$ is an energy-preserving channel. Energy preservation for $\mathcal{M}$ simply follows from the condition $U^\dagger H_{\text{sys}} U = H_{\text{sys}}$. Note that this conclusion is independent of the validity of Yanase's condition, because the sum over all measurement outcomes yields the identity operator, no matter what type of measurement is performed.

The physical model discussed so far motivates the following definition:

Definition 1. We say that a quantum instrument $\{\mathcal{M}_x\}_{x \in \mathsf{X}}$ is *energy-preserving* if the average channel $\mathcal{M} := \sum_{x \in \mathsf{X}} \mathcal{M}_x$ is energy-preserving.

Energy-preserving instruments are exactly the instruments that can be implemented (in principle) at zero energy cost. Precisely, every such instrument can be realized via an energy-preserving interaction, followed by the measurement of a meter observable that commutes with the energy:

Theorem 2. Let $\{\mathcal{M}_x\}_{x \in \mathsf{X}}$ be a quantum instrument transforming states on $\mathcal{H}$. Then, $\{\mathcal{M}_x\}_{x \in \mathsf{X}}$ is energy-preserving if and only if one has

$$\mathcal{M}_x(\rho) = \text{Tr}_{\text{env}}[(I_{\text{sys}} \otimes Q_x)U(\rho \otimes |\phi_0\rangle\langle\phi_0|)U^\dagger] \quad \forall x \in \mathsf{X},$$

where $|\phi_0\rangle$ is a ground state of the environment's Hamiltonian, U satisfies the conditions $U^\dagger H_{\text{sys}} U = H_{\text{sys}}$ and $U^\dagger H_{\text{env}} U = H_{\text{env}}$, and $\{Q_x\}$ is a projective measurement satisfying Yanase's condition $[Q_x, H_{\text{env}}] = 0 \ \forall x \in \mathsf{X}$.

The proof can be found in Appendix B, while a simple illustration of the result is shown in the following example:

Example 1. Consider a system with two energy levels $E_0 = 0$ and $E_1 = \Delta E$, corresponding to the pure states $|0\rangle$ and $|1\rangle$, respectively. Clearly, the von Neumann instrument for the energy measurement—described by the quantum operations $\mathcal{M}_x(\cdot) = |x\rangle\langle x| \cdot |x\rangle\langle x|$, $x = 0, 1$—is energy-preserving. To implement this instrument, one can use as environment two identical copies of the system, choose the initial state $|\phi_0\rangle = |0\rangle|1\rangle$, and engineer a joint evolution U satisfying

$$U|0\rangle|0\rangle|1\rangle = |0\rangle|0\rangle|1\rangle,$$
$$U|1\rangle|0\rangle|1\rangle = |1\rangle|1\rangle|0\rangle,$$
$$U|0\rangle|1\rangle|0\rangle = |0\rangle|1\rangle|0\rangle,$$
$$U|1\rangle|1\rangle|0\rangle = |1\rangle|0\rangle|1\rangle,$$
$$U|0\rangle|0\rangle|0\rangle = |0\rangle|0\rangle|0\rangle,$$
$$U|1\rangle|0\rangle|0\rangle = |1\rangle|0\rangle|0\rangle,$$
$$U|0\rangle|1\rangle|1\rangle = |0\rangle|1\rangle|1\rangle,$$
$$U|1\rangle|1\rangle|1\rangle = |1\rangle|1\rangle|1\rangle.$$

By measuring the meter observable $M = |1\rangle\langle 1| \otimes |0\rangle\langle 0|$ on the environment, one then obtains the instrument $\{\mathcal{M}_0, \mathcal{M}_1\}$ as effective evolution of the system.

It is worth stressing that, despite the fact that the energy is preserved on average, its expectation value can fluctuate due to postselection. For instance, in Example 1 one can decide to postselect the output state $|1\rangle$. With probability $1/2$, the postselection will transform the state $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ into the state $|1\rangle$, whose energy is twice the expected energy of $|+\rangle$. Still, the transformation will take place at no energy cost, because both the interaction and the measurement of the meter observable preserve the energy. A further discussion on this point is provided in Appendix C.

Motivated by the above discussion, we put forward the following:

Definition 2. A probabilistic transformation $\mathcal{M}_0$ is *energy-preserving* iff there exists an energy-preserving instrument $\{\mathcal{M}_x\}_{x \in \mathsf{X}}$ and an outcome $x_0 \in \mathsf{X}$ such that $\mathcal{M}_0 = \mathcal{M}_{x_0}$.

Probabilistic energy-preserving transformations can be demonstrated in engineered quantum systems with a high

degree of control. For example, a proposal for an experimental amplification of weak coherent states via probabilistic energy-preserving transformations was recently reported by Partanen *et al.* in Ref. [65], where high-fidelity amplification was achieved using only passive optical elements.

D. Energy-preserving and covariant instruments: The stationary case

We now show that energy-preserving instruments play a central role in the optimization of probabilistic operations. Consider an instrument whose outcome probabilities are not affected by time translations, namely

$$\text{Tr}[\mathcal{M}_x(U_t \rho U_t^\dagger)] = \text{Tr}[\mathcal{M}_x(\rho)], \quad \forall x \in \mathsf{X}, \forall t \in \mathbb{R}, \quad (23)$$

where ρ is a generic quantum state and $U_t = \exp[-it H_{\text{sys}}/\hbar]$.

We call the instruments satisfying Eq. (23) *stationary*. A common example is that of stationary instruments with outcomes $\mathsf{X} = \{\text{succ,fail}\}$. In this case, the stationarity condition (23) simply means that the probability of implementing the desired transformation $\mathcal{M}_{\text{succ}}$ is the same at every time. Stationary two-outcome instruments can be used to describe tasks like probabilistic covariant cloning and probabilistic amplification, or more generally, any task where the goal is to probabilistically transform a set of states generated by time evolution. An important class of stationary instruments are those that are *covariant* under time evolution, namely

$$\mathcal{M}_x(U_t \rho U_t^\dagger) = V_t \mathcal{M}_x(\rho) V_t^\dagger \quad \forall t \in \mathbb{R}, \quad (24)$$

where ρ is an arbitrary input state and $\{V_t | t \in \mathbb{R}\}$ is the unitary representation of the translation group describing the time evolution of the output system.

Energy-preserving instruments and stationary covariant instruments are closely related. First of all, every energy-preserving instrument is stationary and covariant; indeed, the block form of the Kraus operators (13) implies that every quantum operation $\mathcal{M}_x$ satisfies the covariance condition with

$$U_t = V_t = \exp[-it H_{\text{sys}}/\hbar]. \quad (25)$$

Moreover, energy-preserving instruments are the key probabilistic element at the basis of every stationary covariant instrument:

Proposition 1. Let $\{\mathcal{M}_x\}_{x \in \mathsf{X}}$ be a stationary covariant instrument transforming states on $\mathcal{H}_{\text{in}}$ into states on $\mathcal{H}_{\text{out}}$. Then, $\{\mathcal{M}_x\}_{x \in \mathsf{X}}$ can be decomposed as

$$\mathcal{M}_x = \mathcal{C}_x \mathcal{P}_x \quad \forall x \in \mathsf{X},$$

where $\{\mathcal{P}_x\}_{x \in \mathsf{X}}$ is a pure energy-preserving instrument transforming states on $\mathcal{H}_{\text{in}}$ into states on $\mathcal{H}_{\text{in}}$ and $\mathcal{C}_x$ is a covariant channel transforming states on $\mathcal{H}_{\text{in}}$ into states on $\mathcal{H}_{\text{out}}$.

This result, proven in Appendix D, provides additional motivation to the study of energy-preserving instruments. Indeed, there is a large class of tasks where the optimal probabilistic strategy is described by a stationary covariant instrument—this is the case, e.g., of phase-covariant probabilistic cloning [27–29] and probabilistic amplification [32]. Proposition 1 establishes that energy-preserving instruments are the canonical probabilistic element in all these tasks. The search for the optimal quantum operation is then split into two sub-problems: (i) the search for the optimal energy-preserving

instrument and (ii) the search for the optimal deterministic operation. Now, the optimization of deterministic operations has been studied extensively in the literature, and the solution of problem (ii) is known in a number of relevant cases. In all these cases, the search for the optimal probabilistic operation is reduced to the search of the optimal energy-preserving instrument. A general method for the solution of the problem will be provided in Sec. III.

E. Energy-preserving and covariant instruments: The nonstationary case

The relation between energy-preserving and covariant instruments can also be extended to nonstationary cases where the outcome probabilities are transformed by time evolution as

$$\text{Tr}[\mathcal{M}_x(U_t \rho U_t^\dagger)] = \text{Tr}[\mathcal{M}_{f_{-t}(x)}(\rho)], \quad \forall x \in \mathsf{X}, \forall t \in \mathbb{R}. \quad (26)$$

Here $f_t : \mathsf{X} \to \mathsf{X}$ is a function representing the action of the time evolution on the outcomes. This means that f_t is invertible and satisfies the condition

$$f_{t_1} \circ f_{t_2} = f_{t_1 + t_2}, \quad \forall t_1, t_2 \in \mathbb{R}.$$

Note that stationary instruments are included as a special case, because Eq. (23) can be obtained from Eq. (26) by setting $f_t(x) = x$ for every x and t.

In the nonstationary case, covariant instruments are defined by the condition [63]

$$\mathcal{M}_x(U_t \cdot U_t^\dagger) = V_t \mathcal{M}_{f_{-t}(x)}(\cdot) V_t^\dagger, \quad \forall x \in \mathsf{X}, \forall t \in \mathbb{R}. \quad (27)$$

An example of nonstationary covariant instrument arises in the task of phase estimation, where the time evolution is periodic and the set of outcomes is $\mathsf{X} = [0, 2\pi)$. A covariant measurement of phase satisfies Eq. (26), with

$$f_t(\theta) = \theta + \omega t \mod 2\pi, \quad \forall \theta \in [0, 2\pi), t \in \mathbb{R},$$

where ω is the frequency of the oscillator. Another example arises in probabilistic phase estimation [26], where the set of outcomes is

$$\mathsf{X} = \{0, 2\pi\} \cup \{\text{fail}\},$$

and the outcome "fail" occurs when no phase estimate is produced. In this case, it is natural to require

$$f_t(\text{fail}) = \text{fail}, \quad t \in \mathbb{R},$$

implying that the probability of failure is the same at all times.

For probabilistic phase estimation, the set of outcomes can be partitioned into two orbits: one orbit containing all the outcomes in $[0, 2\pi)$ and one orbit containing the single outcome $\{\text{fail}\}$. In general, for a nonstationary instrument satisfying Eq. (26), the outcome set X can be partitioned into disjoint orbits, as

$$\mathsf{X} = \bigcup_{y \in \mathsf{Y}} \mathsf{O}_y, \quad (28)$$

where each O_y is an orbit under the action of the translation group and the set Y labels the different orbits. With this

notation, every covariant instrument can be decomposed as follows:

Proposition 2. Let $\{\mathcal{M}_x\}_{x\in\mathsf{X}}$ be a covariant instrument transforming states on $\mathcal{H}_{\text{in}}$ into states on $\mathcal{H}_{\text{out}}$. Then, $\{\mathcal{M}_x\}_{x\in\mathsf{X}}$ can be decomposed as

$$\mathcal{M}_x = \mathcal{M}_x^{(y)} \circ \mathcal{P}_y, \tag{29}$$

where y is the orbit label defined in Eq. (28), $\{\mathcal{P}_y\}_{y\in\mathsf{Y}}$ is a pure, energy-preserving instrument transforming states on $\mathcal{H}_{\text{in}}$ into states on $\mathcal{H}_{\text{in}}$, and $\{\mathcal{M}_x^{(y)}\}_{x\in\mathsf{O}_y}$ are covariant instruments transforming states on $\mathcal{H}_{\text{in}}$ into states on $\mathcal{H}_{\text{out}}$.

The proof is provided in Appendix E. Physically, Proposition 2 tells us that we can realize every time-covariant covariant instrument through a preselection implemented with energy-preserving operations, followed by conditional measurements that estimate the action of time translations.

In the case of probabilistic phase estimation, this implies that every probabilistic phase measurement can be broken down into a two-outcome, energy-preserving filter $\{\mathcal{P}_{\text{succ}}, \mathcal{P}_{\text{fail}}\}$, followed by a deterministic phase measurement $\{\mathcal{M}_x^{(\text{succ})}\}_{x\in[0,2\pi)}$ in the successful case. Physically, this means that the postselection on the measurement outcomes can be freely transformed into a preselection on the input state. Since the preselection can be done at zero energy cost, our result shows that the energy cost of every probabilistic phase measurement is equal to the energy cost of a corresponding deterministic measurement.

Proposition 2 has also important implications for the search for the optimal phase estimation strategy with a desired probability of success. The optimization problem is split into two subproblems: (i) the search for the optimal energy-preserving instrument and (ii) the search for the optimal deterministic estimation. Since the latter is known in a number of cases [63], Proposition 2 reduces the optimization to the search for the optimal energy-preserving instrument. It is also important to stress that every point in the trade-off between precision of phase estimation and probability of success can be explored by applying the optimal phase measurement after a probabilistic preselection done at zero energy cost.

III. OPTIMAL ENERGY-PRESERVING OPERATIONS

We are now ready to start the search for the optimal operations that transform a coherent superposition of energy eigenstates into another. In this section we formalize the problem and address the optimality question, providing the general form of the optimal energy-preserving operations.

A. How well can we implement a desired state transition without exchanging energy?

Regarding energy as a resource, a natural question is how well a desired task can be achieved without the assistance of external energy sources. Consider the most basic task: transforming a pure input state $|\varphi\rangle$ into a target output state $|\psi\rangle$. For example, the input could be a weak coherent state with known amplitude but unknown phase, and the target could be another coherent state with the same phase but with larger amplitude. The problem of amplifying laser pulses using energy-preserving operations was recently studied in Ref. [65],

where the authors showed that a nearly perfect amplification can be achieved probabilistically by exploiting the quantum fluctuations of the field, without drawing any energy from the outside. Another interesting example is quantum cloning: Suppose that a spin-1/2 particle, immersed in a magnetic field, is prepared in a superposition of spin up and spin down. As a result, the particle will precess around the direction of the magnetic field and its state will evolve in time. How well can we copy the time information without tapping external energy sources? Note that both in the amplification and in the cloning example, the input and the target states are drawn from a *set* of states: more precisely, the problem is to transform the input state $|\varphi_t\rangle = e^{-itH_{\text{sys}}/\hbar}|\varphi\rangle$ into the target state $|\psi_t\rangle = e^{-itH_{\text{sys}}/\hbar}|\psi\rangle$ for an arbitrary (and possibly unknown) value of the parameter t. However, since we require our operations to be energy-preserving, we do not need to optimize them for every value of t: indeed, every energy-preserving transformation that approximates the transition $|\varphi\rangle \to |\psi\rangle$ will do equally well in approximating the transition $|\varphi_t\rangle \to |\psi_t\rangle$, due to the covariance condition of Eqs. (24) and (25). This point is made clear if we measure the performance of our operations in terms of the fidelity between the target state and the actual output. For a probabilistic transformation $\mathcal{M}_x$, the fidelity between target state and the actual output state is

$$F_{x,t} := \langle\psi_t|\rho_x'(t)|\psi_t\rangle, \quad \rho_x'(t) = \frac{\mathcal{M}_x(|\varphi_t\rangle\langle\varphi_t|)}{\text{Tr}[\mathcal{M}_x(|\varphi_t\rangle\langle\varphi_t|)]}.$$

Using the covariance condition one immediately sees that $F_{x,t}$ is independent of t. Physically, this means that energy-preserving transformations perform equally well on all possible inputs that are connected by time evolution.

More generally, it is interesting to consider the probabilistic transformations obtained by postselection over the outcomes of a quantum measurement. We will call a *filter* an instrument $\{\mathcal{M}_x\}_{x\in\mathsf{X}}$ along with a partition of outcome set X into two disjoint subsets—X_{succ} and X_{fail}—which correspond to successful and unsuccessful instances, respectively. Averaging the fidelity over the successful instances, we obtain the value

$$F = \sum_{x\in\mathsf{X}_{\text{succ}}} p(x|\text{succ})\langle\psi_0|\rho_x'(0)|\psi_0\rangle, \tag{30}$$

where $p(x|\text{succ})$ is the conditional probability of obtaining x given that a successful outcome has occurred. Making the filter explicit, the average fidelity can be rewritten as

$$F = \frac{\langle\psi|\mathcal{M}_{\text{succ}}(|\varphi\rangle\langle\varphi|)|\psi\rangle}{p_{\text{succ}}}, \tag{31}$$

where $\mathcal{M}_{\text{succ}}$ is the quantum operation defined by

$$\mathcal{M}_{\text{succ}} := \sum_{x\in\mathsf{X}_{\text{succ}}} \mathcal{M}_x \tag{32}$$

and p_{succ} is the probability of success

$$\begin{aligned}
p_{\text{succ}} &= \text{Tr}[\mathcal{M}_{\text{succ}}(|\varphi\rangle\langle\varphi|)] \\
&= \text{Tr}[\mathcal{M}_{\text{succ}}^\dagger(I)|\varphi\rangle\langle\varphi|] \\
&= \langle\varphi|P_{\text{succ}}|\varphi\rangle, \\
P_{\text{succ}} &:= \mathcal{M}_{\text{succ}}^\dagger(I). \tag{33}
\end{aligned}$$

In a realistic situation one will be interested not only in maximizing the fidelity but also in having a sufficiently high probability of success. Requiring the success probability to be larger than a given threshold, the problem becomes to find the energy-preserving quantum operation $\mathcal{M}_{\text{succ}}$ that maximizes the fidelity.

B. Deterministic transitions: Optimality of eigenstate alignment

Let us consider first the case $p_{\text{succ}} = 1$. In this case, the optimization problem has a closed-form solution, corresponding to an operation that we call *eigenstate alignment*. Given two superpositions of energy eigenstates, eigenstate alignment is an energy-preserving unitary operation that transforms the eigenstates in one superposition into the eigenstates in the other. Precisely, let us decompose the pure states $|\varphi\rangle$ and $|\psi\rangle$ as

$$|\varphi\rangle = \sum_E \sqrt{p_E} |\varphi_E\rangle \quad \text{and} \quad |\psi\rangle = \sum_E \sqrt{q_E} |\psi_E\rangle \quad (34)$$

with

$$p_E = \|P_E|\varphi\rangle\|^2, \quad q_E = \|P_E|\psi\rangle\|^2,$$
$$|\varphi_E\rangle := \frac{P_E|\varphi\rangle}{\|P_E|\varphi\rangle\|}, \quad |\psi_E\rangle := \frac{P_E|\psi\rangle}{\|P_E|\psi\rangle\|}, \quad (35)$$

P_E being the projector on the eigenspace of H_{sys} with energy E. In the rest of the paper, we will extensively use the notations p_E and q_E for the energy spectrum of the initial and final state, respectively. With this notation, we say that a unitary operator U is an *eigenstate alignment* of $|\varphi\rangle$ with $|\psi\rangle$ if U is energy-preserving and

$$U|\varphi_E\rangle = |\psi_E\rangle \quad \forall E : p_E \neq 0, q_E \neq 0. \quad (36)$$

It is immediate to see that an eigenstate alignment exists for every pair of pure states. In general, eigenstate alignments are not unique, unless the spectrum of H_{sys} is nondegenerate and every energy E in the spectrum satisfies the conditions $p_E \neq 0$ and $q_E \neq 0$. The importance of eigenstate alignment comes from the following:

Theorem 3. For $p_{\text{succ}} = 1$, eigenstate alignment achieves the maximum fidelity for the transition $|\varphi\rangle \to |\psi\rangle$. The maximum fidelity is given by

$$F_{\text{det}} = \left(\sum_E \sqrt{p_E q_E} \right)^2. \quad (37)$$

The proof is provided in Appendix F. Theorem 3 shows that the optimal energy-preserving channel can be chosen to be unitary without loss of generality. In this case, no interaction with the environment is needed. We stress that the optimality of unitary transformations is a specific feature of the energy-preserving framework. Unitary transformations may not be optimal in the broader class of phase-covariant channels—for example, they are sometimes suboptimal for cloning qubit states on the equator of the Bloch sphere [66].

C. Probabilistic transitions: Optimality of pure quantum operations

Let us move to the general case $p_{\text{succ}} \leqslant 1$. We now show that, without loss of generality, the optimal quantum operation

$\mathcal{M}_{\text{succ}}$ can be chosen to be *pure*, i.e., of the form $\mathcal{M}_{\text{succ}}(\cdot) = M_{\text{succ}} \cdot M_{\text{succ}}^\dagger$ for some suitable operator M_{succ}. To prove this result, we provide a construction that transforms any given quantum operation $\mathcal{M}_{\text{succ}}$ into a pure quantum operation $\mathcal{M}'_{\text{succ}}$ with the same probability of success and possibly a higher fidelity. The construction is based on an ingredient that we name the *Lüders reduction*.

1. Lüders reduction

The Lüders reduction transforms a given quantum operation into a pure quantum operation with the same probability of success. Specifically, the Lüders reduction of a quantum operation $\mathcal{M}$ is the pure quantum operation $\mathcal{P}$ defined by

$$\mathcal{P}(\cdot) = \sqrt{P} \cdot \sqrt{P}, \quad P = \mathcal{M}^\dagger(I). \quad (38)$$

When P is a projector, the quantum operation $\mathcal{P}$ coincides with the state transformation defined by Lüders in his treatment of the measurement process [67]. When P is not a projector, $\mathcal{P}$ is often called the "generalized Lüders transformation" associated with P [68].

By construction, a quantum operation and its Lüders reduction have the same probability: For every quantum state ρ one has

$$\text{Tr}[\mathcal{P}(\rho)] = \text{Tr}[\sqrt{P}\rho\sqrt{P}]$$
$$= \text{Tr}[P\rho]$$
$$= \text{Tr}[\mathcal{M}^\dagger(I)\rho]$$
$$= \text{Tr}[\mathcal{M}(\rho)]. \quad (39)$$

Among the quantum operations that happen with the same probability as $\mathcal{M}$, the Lüders reduction can be characterized as the "least noisy," meaning that every other quantum operation can be reproduced by applying a noisy channel to the output of $\mathcal{P}$:

Proposition 3. Every quantum operation $\mathcal{M}$ can be decomposed as $\mathcal{M} = \mathcal{C} \circ \mathcal{P}$ where $\mathcal{P}$ is the Lüders reduction of $\mathcal{M}$ and $\mathcal{C}$ is a suitable quantum channel. Moreover, if $\mathcal{M}$ is energy-preserving, then also $\mathcal{P}$ and $\mathcal{C}$ are energy-preserving.

The proof is provided in Appendix G. Using this result, the search for the optimal quantum operation is split into two different problems: the search for an optimal *pure* operation $\mathcal{P}$ and the search for the optimal *deterministic* operation $\mathcal{C}$. Note that this conclusion applies not only to the problem of transforming pure states, but also to the optimization of transitions involving mixed states.

2. Increasing the fidelity without changing the success probability

Combining the Lüders reduction and eigenstate alignment we can turn every quantum operation $\mathcal{M}_{\text{succ}}$ into a pure quantum operation $\mathcal{M}'_{\text{succ}}$ with the same success probability and a possibly higher fidelity. The idea is simple: by Proposition 3, every energy-preserving quantum operation $\mathcal{M}_{\text{succ}}$ can be decomposed as

$$\mathcal{M}_{\text{succ}} = \mathcal{C} \circ \mathcal{P}_{\text{succ}},$$

where $\mathcal{C}$ is an energy-preserving channel and $\mathcal{P}_{\text{succ}}$ is the Lüders reduction given by

$$\mathcal{P}_{\text{succ}}(\cdot) = \sqrt{P_{\text{succ}}} \cdot \sqrt{P_{\text{succ}}}, \quad P_{\text{succ}} = \mathcal{M}_{\text{succ}}^\dagger(I).$$

When the Lüders reduction takes place, the input state $|\varphi\rangle$ is transformed into the pure state

$$|\varphi'\rangle = \frac{\sqrt{P_{\text{succ}}}|\varphi\rangle}{\|\sqrt{P_{\text{succ}}}|\varphi\rangle\|}. \tag{40}$$

Now, the probability of success depends only on the operator P_{succ}. Fixing P_{succ}, we know that the optimal energy-preserving channel for the transition $|\varphi'\rangle \to |\psi\rangle$ is given by eigenstate alignment (Theorem 3). Hence, the fidelity for the quantum operation $\mathcal{M}_{\text{succ}}$ cannot be larger than the fidelity of the quantum operation

$$\mathcal{M}'_{\text{succ}} = \mathcal{U} \circ \mathcal{P}_{\text{succ}}, \tag{41}$$

where $\mathcal{U}$ is the unitary channel corresponding to the eigenstate alignment of $|\varphi'\rangle$ with $|\psi\rangle$. Note that $\mathcal{M}'_{\text{succ}}$ is energy-preserving, because it is the composition of two energy-preserving operations. Summarizing, we have proven the following:

Proposition 4. For every energy-preserving operation $\mathcal{M}_{\text{succ}}$, the energy-preserving operation $\mathcal{M}'_{\text{succ}}$ defined in Eq. (41) has the same success probability and at least the same fidelity in the implementation of the state transition $|\varphi\rangle \to |\psi\rangle$.

Explicitly, the success probability and the fidelity of $\mathcal{M}'_{\text{succ}}$ are given by

$$p_{\text{succ}} = \sum_E p_E \langle \varphi_E | P_{\text{succ}} | \varphi_E \rangle \tag{42}$$

and

$$F = \left(\sum_E \sqrt{p'_E q_E} \right)^2, \quad p'_E = \frac{p_E \langle \varphi_E | P_{\text{succ}} | \varphi_E \rangle}{p_{\text{succ}}}, \tag{43}$$

where p_E and q_E are the probabilities in the input and output states, as defined in Eq. (35). The above expression of the fidelity follows directly from the application of Theorem 3 to the transition $|\varphi'\rangle \to |\psi\rangle$.

Now, since turning a quantum operation into a pure quantum operation can only increase the fidelity, we proved the following:

Corollary 1 (optimality of pure quantum operations). For every fixed value of the success probability, the energy-preserving operation that maximizes the fidelity can be chosen to be pure without loss of generality.

3. Optimal quantum operations from Lagrangian optimization

The optimization of the fidelity for given success probability can be completed by Lagrangian optimization. Let us define the coefficients

$$x_E := \langle \varphi_E | P_{\text{succ}} | \varphi_E \rangle. \tag{}$$

With this definition, the probability of success (42) and the fidelity (43) can be expressed as

$$p_{\text{succ}} = \sum_E p_E x_E \tag{44}$$

and

$$F = p_{\text{succ}}^{-1} \left(\sum_E \sqrt{x_E p_E q_E} \right)^2, \tag{45}$$

respectively. By Lagrangian optimization, we obtain that the optimal filter has a simple structure: the energy spectrum is partitioned into two disjoint subsets, S_0 and S_1, and the coefficients of the optimal transformation are given by

$$x_E = \begin{cases} 1 & E \in \mathsf{S}_0, \\ \frac{p_{\text{succ}} - p(\mathsf{S}_0)}{1 - q(\mathsf{S}_0)} \frac{q_E}{p_E} & E \in \mathsf{S}_1, \end{cases} \tag{46}$$

where $p(\mathsf{S}_0) := \sum_{E \in \mathsf{S}_0} p_E$ and $q(\mathsf{S}_0) := \sum_{E \in \mathsf{S}_0} q_E$. In other words, the optimal filter is completely determined by the choice of the set S_0. Inserting Eq. (46) into Eq. (45), the maximization of the fidelity is reduced to the maximization of the quantity

$$\Omega[S_0] = \sum_{E \in \mathsf{S}_0} \sqrt{p_E q_E} + \sqrt{[p_{\text{succ}} - p(\mathsf{S}_0)][1 - q(\mathsf{S}_0)]}. \tag{47}$$

Examples of quantum operations of the form (46) can be found in Ref. [26], which focused on the specific problem of phase estimation. More examples will be provided in Sec. V.

D. The ultimate limits of probabilistic energy-preserving processes

So far we considered the optimization of the fidelity for fixed success probability. We now remove the constraint and focus only on the maximization of the fidelity. The problem is interesting because it highlights the quantum limits to what is logically possible, no matter how small the probability [69].

The ultimate limit for energy-preserving operations is characterized by the following:

Proposition 5. Let $|\varphi\rangle$ and $|\psi\rangle$ be two generic pure states of a finite-dimensional quantum system. For the transition $|\varphi\rangle \to |\psi\rangle$, the maximum of the fidelity over all energy-preserving quantum operations is

$$F_{\max} = \sum_{E \in \mathsf{Sp}(\varphi) \cap \mathsf{Sp}(\psi)} q_E, \tag{48}$$

where q_E is the probability defined in Eq. (35) and $\mathsf{Sp}(\chi)$ denotes the energy spectrum of a generic state $|\chi\rangle$, defined as

$$\mathsf{Sp}(\chi) := \{E | \langle \chi | P_E | \chi \rangle \neq 0\}.$$

For a quantum operation achieving fidelity $F_{\max}$ the maximum probability of success is given by

$$p_{\text{succ}}^{\max} = \left(\min_{E \in \mathsf{Sp}(\varphi) \cap \mathsf{Sp}(\psi)} \frac{p_E}{q_E} \right) F_{\max}, \tag{49}$$

where p_E is the probability defined in Eq. (35). The quantum operation achieving maximum fidelity with maximum probability is pure and its Kraus operator satisfies the condition

$$M|\varphi_E\rangle = \left[\min_{E' \in \mathsf{Sp}(\varphi) \cap \mathsf{Sp}(\psi)} \sqrt{\frac{p_{E'}}{q_{E'}}} \right] \sqrt{\frac{q_E}{p_E}} |\psi_E\rangle \tag{50}$$

for every energy E in $\mathsf{Sp}(\varphi)$.

The proof is provided in Appendix H.

IV. MULTIROUND RECURSIVE PROTOCOLS

In the previous section we provided a recipe to find the protocol that achieves maximum fidelity for a fixed value of

the success probability. By definition, the resulting protocol is tailor-made to that specific value of the probability. However, in many situations it is useful to have a more flexible protocol, where the experimenter can make successive attempts at realizing the desired transformation and is free to decide on the fly when to stop. In this section we analyze protocols of this form, which we refer to as *recursive protocols*. Under the energy-preserving constraint, we identify the protocol that produces the best possible approximation of the target state at each step. It is important to stress that the protocol does not require an actual experimenter to read the outcomes and to make decisions: in principle, all the measurements and conditional operations can be implemented by a fully quantum machine operating in an energy-preserving fashion.

A. The optimal recursive protocol

Given a sequence of K binary filters with outcomes {succ,fail}, consider the protocol defined by the following instructions:

(1) Set $k = 1$.

(2) If $k \leqslant K$, then apply the kth filter; else terminate.

(3) If the outcome is $x = $ succ, then terminate.

(4) If the outcome is $x = $ fail, then replace k with $k + 1$ and go back to instruction (2).

Recursive protocols of this form are an example of "quantum loop programs," studied in Ref. [70]. All these protocols can be can be visualized as decision trees of the following form:
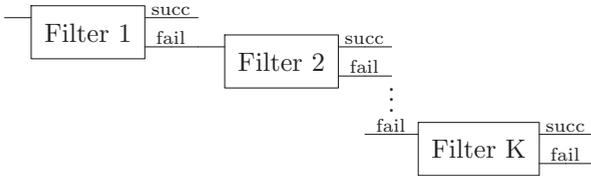
Protocols of this form have been employed for different purposes, including entanglement concentration [71], implementation of quantum gates [72–74], and ancilla-driven computation [75]. One such protocol that is particularly similar to ours is *quantum rejection sampling* [76]. There, the goal is to generate a target superposition $|\psi\rangle$ using a black box U_φ that prepares another superposition $|\varphi\rangle$ from a fixed state $|0\rangle$. The difference between rejection sampling and our problem is that in our case we do not have the black box U_φ at disposal. Instead, we have the coherent superposition $|\varphi\rangle$, which is a strictly weaker resource than the gate U_φ, due to Nielsen and Chuang's no-programming theorem [77].

In our case, the goal of the protocol is to transform a coherent superposition of energy eigenstates into another. Of course, at each step there will be a trade-off between the fidelity with the target and the probability of success. In the simplest scenario, the protocol can be designed to attain the absolute maximum of the fidelity at each round, and to do so with maximum probability of success. An experimenter following such a protocol will have the guarantee that the best possible performance is achieved in each individual round.

We consider the transition $|\varphi\rangle \to |\psi\rangle$ in the case of states $|\varphi\rangle$ and $|\psi\rangle$ in a finite-dimensional Hilbert space $\mathcal{H} \simeq \mathbb{C}^d, d < \infty$, or, more generally, states whose energy spectra intersect

on a finite set of points, with $|\mathsf{Sp}(\varphi) \cap \mathsf{Sp}(\psi)| \leqslant d$. For the optimal protocol we make a list of desiderata in decreasing order of priority: for every $k \in \{1, \ldots, K-1\}$

(1) at the kth round, the successful quantum operation should transform the input state $\rho^{(k)}$ into the target $|\psi\rangle$ with maximum fidelity;

(2) the optimal transition $\rho^{(k)} \to |\psi\rangle\langle\psi|$ must be achieved with maximum probability of success;

(3) the unsuccessful quantum operation at the kth round should produce the state $\rho^{(k+1)}$ that leads to maximum fidelity for the transformation $\rho^{(k+1)} \to |\psi\rangle\langle\psi|$ at the $(k + 1)$th round;

(4) at the final round $(k = K)$ the successful quantum operation should achieve maximum fidelity with maximum probability and, conditional on the fulfillment of this requirement, the unsuccessful quantum operation should achieve maximum fidelity.

The derivation of the optimal protocol is rather technical and is provided in Appendix I. In the following we present the final result of the optimization and discuss the implications of our findings.

At the kth round, the optimal binary filter consists of two pure quantum operations, $\mathcal{B}_{\text{succ}}^{(k)}(\cdot) = B_{\text{succ}}^{(k)} \cdot B_{\text{succ}}^{(k)\dagger}$ and $\mathcal{B}_{\text{fail}}^{(k)}(\cdot) = B_{\text{fail}}^{(k)} \cdot B_{\text{fail}}^{(k)\dagger}$. Since all quantum operations are pure, the state of the system is pure at every round. The input state at the kth round, denoted by $|\varphi^{(k)}\rangle$, can be expanded as

$$|\varphi^{(k)}\rangle = \sum_E \sqrt{p_E^{(k)}}|\varphi_E\rangle,$$

where the energy eigenstates are the same as in the decomposition of $|\varphi\rangle$; cf. Eq. (35). With this notation, the successful quantum operation is determined in an essentially unique way by the condition

$$B_{\text{succ}}^{(k)}|\varphi_E\rangle = \left[\min_{E' \in \mathsf{Sp}(\varphi^{(k)}) \cap \mathsf{Sp}(\psi)} \sqrt{\frac{p_{E'}^{(k)}}{q_{E'}}}\right] \sqrt{\frac{q_E}{p_E^{(k)}}}|\psi_E\rangle$$

$$\forall E \in \mathsf{Sp}(\varphi^{(k)}). \tag{51}$$

Here the only freedom is in the definition of the operator in the subspace spanned by energy eigenstates outside the spectrum of $|\varphi^{(k)}\rangle$. The form of Eq. (51) follows directly from the requirements 1 and 2 in our list of desiderata (cf. Proposition 5). Similarly, the unsuccessful quantum operation is determined in an essentially unique way by the condition

$$B_{\text{fail}}^{(k)} = \sqrt{I - B_{\text{succ}}^{(k)\dagger}B_{\text{succ}}^{(k)}}. \tag{52}$$

The form of Eq. (52) follows from the requirement 3 in our list. Remarkably, the quantum operation $\mathcal{B}_{\text{fail}}^{(k)}$ does not maximize only the fidelity achievable from the input state $\rho^{(k+1)}$, but also the *probability* that maximum fidelity is achieved.

B. Fidelity and success probability

The optimal protocol is specified recursively by Eqs. (51) and (52). Making the dependence on the input and target states explicit, it is possible to derive closed formulas for the fidelity and the success probability. To this purpose, we need to introduce some notation. Given a pair of pure states $|\varphi\rangle$ and $|\psi\rangle$ and given the corresponding probabilities p_E and q_E

defined in Eq. (35), we define the ratio p_E/q_E for all values of the energy in $\mathsf{Sp}(\varphi) \cap \mathsf{Sp}(\psi)$. Then, we arrange the values of the ration r_E in increasing order as

$$0 < r_1 < r_2 < \cdots < r_L,$$

where r_L is the maximum ratio. Clearly, by the assumption of finite dimensionality, L satisfies the relation

$$L \leqslant |\mathsf{Sp}(\varphi) \cap \mathsf{Sp}(\psi)| \leqslant d.$$

For every possible value r_i, we consider the set of energy eigenvalues R_i defined as

$$\mathsf{R}_i := \left\{ E \in \mathsf{Sp}(\varphi) \cap \mathsf{Sp}(\psi) \middle| \frac{p_E}{q_E} = r_i \right\} \tag{53}$$

and we denote the union of the first k sets by

$$\mathsf{U}_k := \bigcup_{i=1}^{k} \mathsf{R}_i. \tag{54}$$

With this definition, the fidelity and the success probability at the kth step can be expressed as

$$F_{\max}^{(k)} = \sum_{E \in \mathsf{Sp}(\varphi) \backslash \mathsf{U}_{k-1}} q_E \tag{55}$$

and

$$p_{\text{succ}}^{(k)} = (r_k - r_{k-1}) \cdot F_{\max}^{(k)}, \tag{56}$$

respectively. The proof is presented in Appendix J. Note that the fidelity is strictly decreasing with k, reaching zero for $k = L$. In other words, it is useless to consider protocols with more than L rounds.

The explicit expressions given by Eqs. (55) and (56) turn out to be very useful for studying the trade-off between fidelity and success probability.

Indeed, they allow us to evaluate the probability that the protocol succeeds in one of the first T rounds, given by

$$p_{\text{succ}}(T) := \sum_{k=1}^{T} p_{\text{succ}}^{(k)}$$
$$= \sum_{E \in \mathsf{U}_{T-1}} p_E + r_T F_{\max}^{(T)}, \tag{57}$$

and to observe its scaling with the average fidelity achieved in the first T steps, given by

$$F(T) := \frac{\sum_{k=1}^{T} p_{\text{succ}}^{(k)} F_{\max}^{(k)}}{p_{\text{succ}}(T)}. \tag{58}$$

The trade-off curve between $F(T)$ and $p_{\text{succ}}(T)$ will be illustrated in Sec. V for a number of concrete examples.

C. Output states and termination time of the protocol

In addition to the fidelity and success probability, it is useful to know what states are produced at every step of the protocol. Assuming that the total number of rounds is upper bounded as $K \leqslant L$, the explicit expression of the output state produced at the kth round can be obtained as follows. We regard the recursive protocol as a quantum instrument, with outcomes in

the set $\{1, \ldots, K + 1\}$. The outcome k corresponds to the pure quantum operation with Kraus operator

$$M_k := \begin{cases} B_{\text{succ}}^{(k)} B_{\text{fail}}^{(k-1)} \cdots B_{\text{fail}}^{(1)} & k = 1, \ldots, K, \\ B_{\text{fail}}^{(K)} B_{\text{fail}}^{(K-1)} \cdots B_{\text{fail}}^{(1)} & k = K + 1. \end{cases} \tag{59}$$

For $k \leqslant K$, the Kraus operators are characterized in Appendix J. Using this characterization, we show that the output state in the case of success at the kth round is given by

$$|\psi^{(k)}\rangle := \frac{M_k |\varphi\rangle}{\|M_k |\varphi\rangle\|}$$
$$= \frac{1}{\sqrt{F_{\max}^{(k)}}} \sum_{E \in \mathsf{Sp}(\varphi) \backslash \mathsf{U}_{k-1}} \sqrt{q_E} |\psi_E\rangle. \tag{60}$$

Note that $|\psi^{(k)}\rangle$ is a truncated version of the target state, with the energy spectrum deprived of all the values in U_{k-1} and of all the values that are not in in the spectrum of $|\varphi\rangle$. The energy spectrum of the output state is eroded from one step to the next: each iteration of the protocol produces a state with a strictly lesser amount of coherence in the energy eigenbasis. Due to the assumption of finite dimensionality, the process of erosion terminates in a finite number of steps, equal to $T_{\max}$. Protocols with more than $T_{\max}$ rounds terminate after $T_{\max}$ steps, meaning that the probability of success satisfies

$$p_{\text{succ}}(T) = 1 \quad \forall\, T > T_{\max}.$$

The fact that the protocol is guaranteed to terminate in a finite time is an appealing feature. It is worth stressing that the termination time $T_{\max}$ is upper bounded by the number of distinct energy levels of the system, which can be much smaller than the dimension of the Hilbert space, as in the following:

Example 2. Consider the case of N identical noninteracting systems of dimension d. In this case the total Hamiltonian is the sum of the single-system Hamiltonians, and its number of energy levels is upper bounded by the number of partitions of N into d non-negative numbers (see, e.g., [27]). We then have that the number of rounds needed to terminate is upper bounded as

$$T_{\max} \leqslant \binom{d + N - 1}{N} < (N + 1)^{d-1},$$

i.e., by a polynomial in N. Even if the probability of success in the first round is exponentially small in N, as in the case of quantum super-replication [27,28], the recursive protocol is guaranteed to reach unit probability in a polynomial number of iterations.

D. Increasing the fidelity of the recursive protocol

At every iteration of the recursive protocol, the total probability of success increases, while the average fidelity decreases. In general, the relation between fidelity and probability of success is not optimal, because the histories leading to successful outcomes are mixed incoherently: at the Tth step, the successful quantum operation has the form

$$\mathcal{M}^{(T)}(\rho) = \sum_{k=1}^{T} M_k \rho M_k^\dagger, \tag{61}$$

where M_k are the Kraus operators defined in Eq. (59). Now, we have a systematic method to increase the fidelity while keeping the same success probability: the method is to take the Lüders reduction of $\mathcal{M}^{(T)}$ and to perform eigenstate alignment on the output. The following paragraphs highlight the main features of this method.

1. Coherent coarse-graining

The Lüders reduction transforms the quantum operation $\mathcal{M}^{(T)}$ into the pure quantum operation

$$\mathcal{P}^{(T)}(\cdot) = \sqrt{P^{(T)}}(\cdot)\sqrt{P^{(T)}}$$

with

$$P^{(T)} = \mathcal{M}^{(T)\dagger}(I) = \sum_{k=1}^{T} M_k^\dagger M_k.$$

The technique of joining different quantum operations into a single one will be an important tool in the following. For this reason, it is convenient to have a name for it:

Definition 3. We call $\mathcal{P}^{(T)}$ the *coherent coarse-graining* of the quantum operations $\{\mathcal{M}_k | k = 1, \ldots, T\}$.

An intuitive way to visualize coherent coarse-graining is through a generalization of the double-slit experiment. Consider an interference experiment involving multiple slits [78]. When detectors are placed on the slits, the passage of a particle through the kth slit will trigger the occurrence of the quantum operation $\mathcal{M}_k$. When the detectors at the first T slits are removed, the passage through these slits will result into the coherently coarse-grained operation $\mathcal{P}^{(T)}$.

Note that the flexibility of the recursive protocol is lost after coherent coarse-graining: when multiple histories are coherently combined, it is not possible anymore to choose on the fly when to stop the protocol. Still, the advantage of coherent coarse-graining is that it provides a heuristic way to construct lower bounds on the probability-fidelity trade-off.

2. Eigenstate alignment

By construction, coherent coarse-graining does not change the probability of success. The fidelity is then increased by eigenstate alignment, achieved by any energy-preserving unitary U such that

$$U|\varphi_E\rangle = |\psi_E\rangle \quad \forall E \in \mathsf{Sp}(\varphi) \cap \mathsf{Sp}(\psi).$$

Note that the operation of eigenstate alignment does not depend on how many rounds of the protocols are coarse-grained. The operation could be performed even before the filter is applied, provided that one suitably adapts the definition of the filter.

When combined, coherent coarse-graining and eigenstate alignment yield the pure quantum operation

$$\mathcal{M}^{(T)\prime}(\cdot) = M^{(T)\prime} \cdot M^{(T)\prime\dagger}, \quad M^{(T)\prime} := U\sqrt{\sum_{k=1}^{T} M_k^\dagger M_k}, \quad (62)$$

whose action on the energy eigenstates is given by

$$M^{(T)\prime}|\varphi_E\rangle = \begin{cases} |\psi_E\rangle & E \in \mathsf{U}_T, \\ \sqrt{r_T \frac{q_E}{p_E}}|\psi_E\rangle & E \notin \mathsf{Sp}(\varphi) \setminus \mathsf{U}_T, \end{cases} \quad (63)$$

[see Eq. (J8) of Appendix J for the explicit derivation]. For T larger than the termination time $T_{\max}$, our construction eventually yields the *optimal* energy-preserving channel for the transition $|\varphi\rangle \to |\psi\rangle$ (cf. Theorem 3).

3. The performance of the coherently coarse-grained protocol

Since taking eigenstate alignment as an obliged step in the optimal operation, we refer to the quantum operation $\mathcal{M}^{(T)\prime}$ simply as a *coherent coarse-graining* (of the first T steps of the protocol). By construction, the probability of success of the quantum operation $\mathcal{M}^{(T)\prime}$ is equal to the probability that the original (non-coarse-grained) protocol, succeeds within T steps [cf. Eq. (56)]. On the other hand, the fidelity can be evaluated explicitly by using Eq. (63), which yields

$$F'(T) = \frac{\left[\sum_{E \in \mathsf{U}_T} \sqrt{p_E q_E} + \sqrt{r_T}\sum_{E \in \mathsf{Sp}(\varphi)\setminus\mathsf{U}_T} q_E\right]^2}{\sum_{E \in \mathsf{U}_T} p_E + r_T \sum_{E \in \mathsf{Sp}(\varphi)\setminus\mathsf{U}_T} q_E}. \quad (64)$$

Performing the operation of coherent coarse-graining for different values of T one can obtain a sequence of filters that approximate the optimal curve of the fidelity-probability trade-off. The improvement due to coherent coarse-graining will be illustrated in the next section with a number of concrete examples.

V. APPLICATIONS

In this section we apply the recursive protocol and the method of coherent coarse-graining to the tasks of phase estimation, cloning of quantum clocks, phase-insensitive amplification of coherent states, and approximate correction in ancilla-driven quantum computation.

A. Quantum metrology with probabilistic energy-preserving operations

Here we apply the recursive protocol to the task of phase estimation [63,79]. The main idea is the following: When the phase is encoded in a quantum state in a suboptimal way, one can try to improve the precision of phase estimation by first transforming the state into the optimal state. Of course, such transition cannot take place deterministically—for otherwise the original state would have been already optimal. However, a probabilistic protocol can produce good approximations of the optimal state and, conditionally on the success of the probabilistic transformations, it can enable an improved phase estimation. In the following we will use our recursive protocol to scan the trade-off curve between fidelity and probability of success.

To illustrate the idea, we consider the simple case where the phase is encoded into the state of a two-level quantum system, as

$$|\varphi_\theta\rangle = e^{-i\theta Z}|\varphi\rangle \quad \theta \in [0, 2\pi),$$

with $Z = |0\rangle\langle 0| - |1\rangle\langle 1|$ and $|\varphi\rangle = (|0\rangle + e^{i\theta}|1\rangle)/\sqrt{2}$. We assume that N identical copies of the state are available and search for the optimal strategy to estimate θ. To quantify the precision, we use the gain function $G(\theta, \hat{\theta})$ defined by [63]

$$G(\theta, \hat{\theta}) := \frac{1 + \cos(\theta - \hat{\theta})}{2}.$$

Note that the gain function assigns a larger gain when the estimate $\hat\theta$ is closer to the true value θ, attaining the maximum value 1 if and only if $\hat\theta = \theta$. Then, the goal is to find the estimation strategy that maximizes the average gain

$$\langle G \rangle := \int \frac{d\theta}{2\pi} \int \frac{d\hat\theta}{2\pi} G(\theta,\hat\theta) \langle \varphi_\theta |^{\otimes N} E_{\hat\theta} | \varphi_\theta \rangle^{\otimes N}, \qquad (65)$$

where $\{E_{\hat\theta}\}$ is the positive operator valued measure (POVM) describing the estimation strategy.

For phase estimation with pure states, the optimal POVM has been derived by Holevo [63]. Specifically, for pure states of the form

$$|\Phi_\theta\rangle = \sum_{n=0}^{N} c_n e^{-i\theta n} |n\rangle, \quad c_n \geqslant 0, \forall n \in [0,N], \qquad (66)$$

Holevo's POVM yields the gain

$$\langle G_{\text{det}} \rangle = \tfrac{1}{2} + \tfrac{1}{2} \langle \Phi_0 | \Delta | \Phi_0 \rangle,$$

$$\Delta_{mn} = \tfrac{1}{2} [\delta_{m(n-1)} + \delta_{m(n+1)}]. \qquad (67)$$

In our case, the above expression yields the value

$$\langle G_{\text{det}} \rangle = \frac{1}{2} + \frac{1}{2^{N+1}} \sum_{n=0}^{N-1} \sqrt{\binom{N}{n}\binom{N}{n+1}}$$

$$= 1 - O\left(\frac{1}{N}\right). \qquad (68)$$

Now, when the unknown phase shift $e^{-i\theta Z}$ is probed N times, one can obtain a much better estimate by preparing the optimal input state, which in this case is the "sine state" [80,81]

$$|\varphi_{\text{opt},\theta}\rangle = \sqrt{\frac{2}{N+1}} \sum_{n=0}^{N} \sin\left(\frac{n\pi}{N+1}\right) e^{i\theta n} |n\rangle. \qquad (69)$$

This state achieves the Heisenberg scaling $\langle G \rangle = 1 - O(1/N^2)$. In the following, we will use our recursive protocol to transform the state $|\varphi_\theta\rangle^{\otimes N}$ into approximations of the optimal state $|\varphi_{\text{opt},\theta}\rangle$, which will then be used for state estimation.

Note that the output states of our protocol are of the form (69) at every step. Thanks to this fact, we can apply Holevo's recipe (67) to compute the optimal gain. Precisely, the gain at the kth step is given by

$$\langle G^{(k)} \rangle = \tfrac{1}{2} + \tfrac{1}{2} \langle \psi^{(k)} | \Delta | \psi^{(k)} \rangle, \qquad (70)$$

where $|\psi^{(k)}\rangle$ is the output state at the kth step, given by Eq. (60). Averaging over the first T steps we obtain the gain

$$\langle G_T \rangle := \frac{\sum_{k=1}^{T} p_{\text{succ}}^{(k)} \langle G^{(k)} \rangle}{p_{\text{succ}}(T)}, \qquad (71)$$

where $p_{\text{succ}}^{(k)}$ is the probability of achieving success at the kth step and $p_{\text{succ}}(T) = \sum_{k=1}^{T} p_{\text{succ}}^{(k)}$. The value of the gain can be explicitly calculated using Eqs. (57), (60), and (J8). In Fig. 2 we show the estimation gain for $N = 30$ copies of the input state and for $K = 27$ iterations of the recursive protocol.

The performance of the recursive protocol can be compared with the performances of its coherent coarse-graining. By
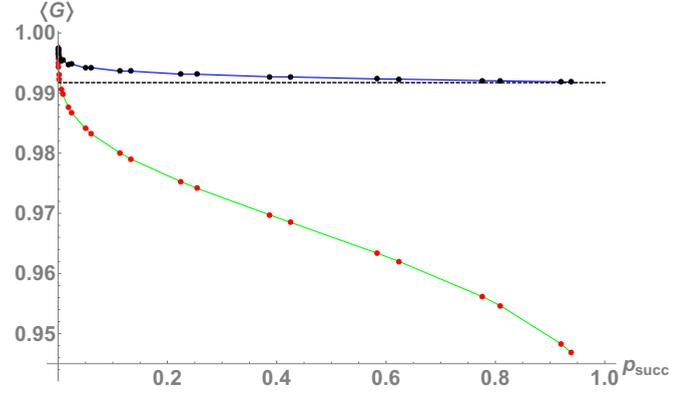
FIG. 2. Probabilistic phase estimation via the recursive protocol and its coherent coarse-graining. The figure shows the trade-off between success probability and average gain for phase estimation with the qubit state $|\varphi_\theta\rangle^{\otimes N}$ for $N = 30$. The green solid line (with numerics represented by red dots) shows the trade-off between estimation gain and success probability for a recursive protocol with $K = 27$ rounds, with the Tth point corresponding to the first T steps. The blue solid line (with numerics represented by the black dots) shows the trade-off for filters generated by coherent coarse-graining, with the Tth point corresponding to the coherent coarse-graining of the first T steps. Note that the gain for the coherent coarse-graining remains higher than the optimal deterministic estimation's gain (the black dashed line) even when the protocol becomes "almost deterministic" (i.e., the probability of success tends to one), while the gain for the recursive protocol drops down quickly with the growth of the success probability.

coherently coarse-graining over the first T rounds, we obtain the average gain given by

$$\langle G'_T \rangle = \tfrac{1}{2} + \tfrac{1}{2} \langle \psi'(T) | \Delta | \psi'(T) \rangle, \qquad (72)$$

with

$$|\psi'(T)\rangle = \frac{M'^{(T)} |\varphi\rangle^{\otimes N}}{\| M'^{(T)} |\varphi\rangle^{\otimes N} \|}$$

and $M^{(T)}$ as in Eq. (62). The estimation gain of the coherent coarse-graining is plotted in Fig. 2. In addition, Fig. 3 shows the scaling of the gain and the success probability with the number of copies N.

B. Converting coherence into metrological precision

In the previous subsection we analyzed the problem of phase estimation with N identical qubits. Here we will consider a one-shot scenario, where the phase has to be estimated from a single copy of the state

$$|\varphi_\theta\rangle := \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} e^{-in\theta} |n\rangle, \qquad (73)$$

consisting of a uniform superposition of the eigenstates of the Hamiltonian $H = \sum_{n=1}^{N-1} n |n\rangle\langle n|$. The above state is the *maximally coherent state* [34] in the energy eigenbasis; that is, it is the most valuable state in the resource theory of coherence. Interestingly, it is *not* the optimal state for phase estimation. Indeed, the estimation gain for the maximally coherent state

(a)

(b)
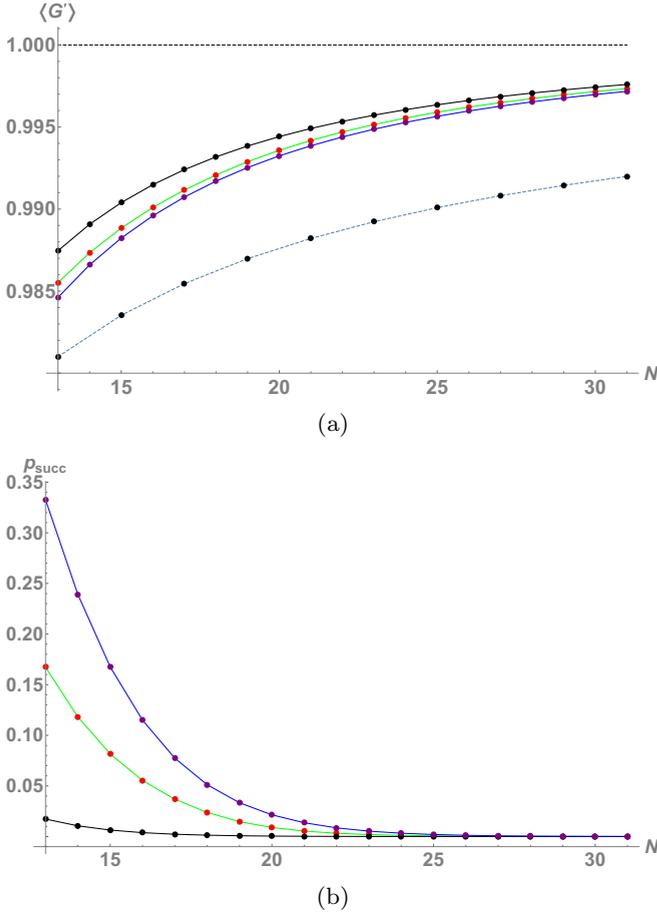
FIG. 3. Scaling of the gain and success probability for coherently coarse-grained protocols. Panel (a) shows the estimation gain as a function of the number of copies N, for three coherently coarse-grained protocols corresponding to different values of T, including $T = 1$ (black line with black dots), $T = 3$ (green line with red dots), and $T = 5$ (blue line with purple dots). The dashed line with black dots represents the optimal deterministic gain $\langle G_{\rm det} \rangle$. Panel (b) shows the decrease of the total success probability as a function of N for different values of T, including $T = 1$ (black line with black dots), $T = 3$ (green line with red dots), and $T = 5$ (blue line with purple dots).

can be evaluated with Eq. (67), which in this case yields

$$\langle G_{\rm det} \rangle = 1 - \frac{1}{2N}. \tag{74}$$

When the number of copies is asymptotically large, the gain approaches its maximum value with the standard quantum limit scaling $1/N$, rather than the Heisenberg scaling $1/N^2$.

We now explore how the maximally coherent state can be transformed into approximations of the optimal state for phase estimation. The performance of the recursive protocol and of its coherent coarse-graining can be evaluated using Eqs. (71) and (72). When the number of iterations T is small compared to the number of energy levels N, the average gain has the simple analytical expression

$$\langle G_T \rangle = 1 - \frac{\pi^2}{2N^2} \left[T(T-1) + \frac{1}{2} \right] + O\left[\left(\frac{T}{N} \right)^3 \right]. \tag{75}$$
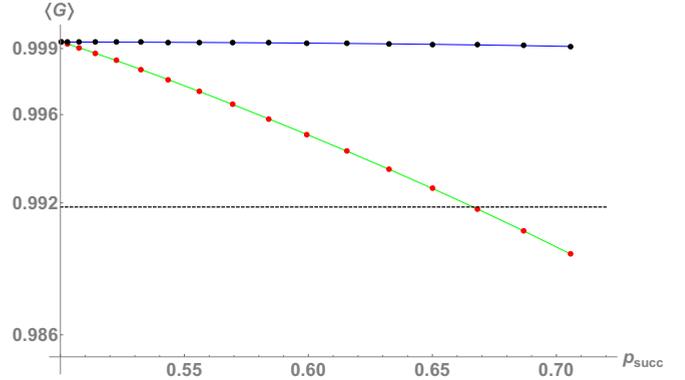
FIG. 4. Probabilistic phase estimation via the recursive protocol and its coherent coarse-graining. The figure shows the trade-off between success probability and average gain for phase estimation with maximally coherent states with $N = 61$. The green solid line (with numerics represented by red dots) shows the probability-gain trade-off for $K = 17$ rounds of the recursive protocol. At the first round the protocol reaches the maximum possible gain, equal to $G_{\rm max} = 99.9\%$, in agreement with the analytical expression of Eq. (75). The blue solid line (with numerics represented by the black dots) shows the trade-off for filters generated by coherent coarse-graining, with the Tth point corresponding to the coherent coarse-graining of the first T steps of the recursive protocol. For the first $K = 17$ rounds the estimation gain of coherent coarse-graining remains approximately equal to $G_{\rm max} = 99.9\%$, although eventually it is bound to decrease to the optimal deterministic value $\langle G_{\rm det} \rangle = 99.2\%$ (black dashed line).

Note that the gain exhibits Heisenberg scaling with the number of energy levels N, with a constant that grows quadratically with the number of rounds T. The success probability can also be evaluated analytically in the regime $N \gg T$ and its value is given by

$$p_{\rm succ}(T) = \frac{1}{2} + \frac{\pi^2}{N^2} \left[T(T-1) + \frac{1}{8} \right] + O\left[\left(\frac{T}{N} \right)^3 \right]. \tag{76}$$

From the above expressions, one can clearly see the trade-off between gain and success probability, which can be made explicit in the trade-off curve

$$\langle G_T \rangle = 1 - \frac{3\pi^2}{16N^2} - \frac{p_{\rm succ}(T) - 1/2}{2}, \quad N \gg T. \tag{77}$$

In Fig. 4 we illustrate the trade-off between the probability of success and the average gain for $N = 61$. The recursive protocol manages to increase the probability of success by approximately 30% from the first round to the 14th, while keeping the average gain above the deterministic gain. In Fig. 5 we show the scaling of the gain and the success probability with the dimension N for different values of T.

Let us evaluate now the performance of coherent coarse-graining. In the $N \gg T$ regime, the gain has the analytical expression

$$\langle G'_T \rangle = 1 - \frac{\pi^2}{4N^2} \left\{ 1 + 4 \left[p_{\rm succ}(T) - \frac{1}{2} \right]^2 \right\}.$$

The trade-off between estimation gain and probability of success is illustrated in Fig. 4 for $N = 61$. Also in this case,
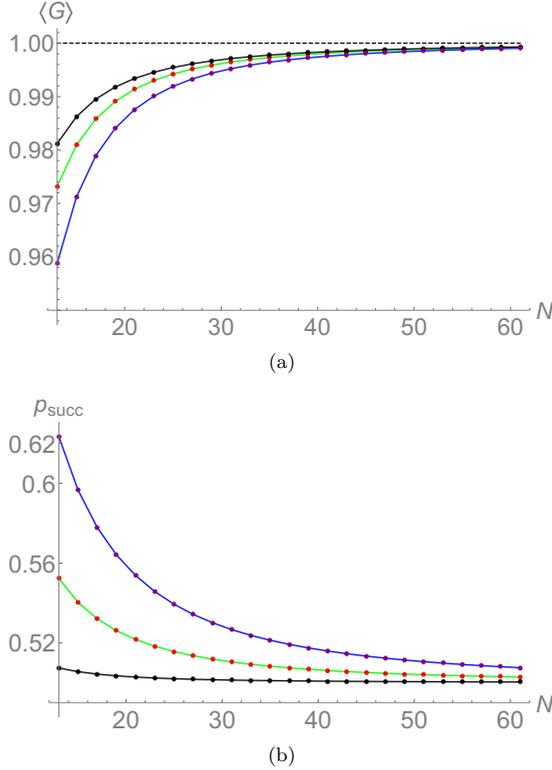
FIG. 5. Scaling of the estimation gain and success probability for the recursive protocol. Panels (a) and (b) show the average gain G and the total success probability (b) as a function of the energy scale N for different values of T, including $T = 1$ (black line with black dots), $T = 2$ (green line with red dots), and $T = 3$ (blue line with purple dots).

one can observe that coherent coarse-graining offers a better trade-off curve than the recursive protocol.

Finally, it is interesting to compare the performance of the coherent coarse-graining with the optimal trade-off curve between gain and probability of success, which is known explicitly in the case under consideration [26]. Remarkably, the comparison shows that for large N the coherent coarse-graining yields exactly the optimal estimation strategy of Ref. [26]. In other words, in this case the coherent coarse-graining of our recursive protocol is asymptotically optimal. At this point, a natural question is whether coherent coarse-graining *always* gives the optimal fidelity/probability trade-off. The answer turns out to be negative: by evaluating Eq. (72) for small values of N (e.g., $N = 10$) we find out that the average gain of the coherent coarse-graining sometimes falls below the threshold of the optimal deterministic gain in Eq. (74), clearly indicating suboptimality in the nonasymptotic regime.

C. Energy-preserving cloning of quantum coherence

Here we consider the problem of quantum cloning [82,83], where the task is to transform N identical copies of an unknown quantum state into a larger number $M \geqslant N$ of approximate copies.

In most cases, the problem has been addressed without imposing any constraint on the cloning process, except for its compatibility with the laws of quantum mechanics. Instead, here we consider copy machines that have to work without any supply of energy for the outside. Consider for example a scenario where one wants to clone the state of a quantum clock [27], given by

$$|\psi_t\rangle = e^{-itH/\hbar}|\psi\rangle,$$

where $H = H^\dagger$ is a suitable Hamiltonian. Here the time parameter t is assumed to be unknown and the copy machine is required to work equally well for every value of t. In order to produce copies without requiring energy from the outside, the machine has to process the N input clocks jointly with a state of $M - N$ "blank clocks," which provide no information about time, but possess sufficient energy to enable the desired transition. Indeed, in order to approximate M perfect copies of the state $|\psi_t\rangle$ the machine should at least be able to produce output states that have energy close to $M\langle\psi|H|\psi\rangle$, meaning that the blank clocks should have energy close to $(M - N)\langle\psi|H|\psi\rangle$. The problem of energy-preserving cloning of clock states is equivalent to the problem of cloning coherence: denoting by $|\beta\rangle$ the blank state, the cloning machine attempts at converting the state $|\psi\rangle^{\otimes N} \otimes |\beta\rangle$ into the state $|\psi\rangle^{\otimes M}$. Choosing the blank state to be an eigenstate of the energy, we have that maximizing the fidelity for the transition $|\psi\rangle^{\otimes N} \otimes |\beta\rangle \to |\psi\rangle^{\otimes M}$ under the energy-preserving restriction is equivalent to maximizing the fidelity of cloning for every instant of time.

In the following we analyze in detail the simplest example of energy-preserving cloning of quantum coherence: we consider N two-level systems, each of them with Hamiltonian $H = \frac{\hbar\omega}{2}Z$ and initially prepared in the coherent superposition $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. The question is how well one can produce $M > N$ approximate copies without paying an energy cost. For simplicity, we assume that the difference $M - N$ is even: under this assumption, we can choose the blank state to be an energy eigenstate with energy *exactly* equal to zero. Specifically, we choose the state $|\beta\rangle = |M - N, 0\rangle$, belonging to the symmetric eigenbasis

$$|L, m\rangle := \frac{\sum_{\pi \in S_L} U_\pi |0\rangle^{\otimes(L+m)/2} \otimes |1\rangle^{\otimes(L-m)/2}}{\sqrt{L![(L+m)/2]![(L-m)/2]!}},$$

where S_L denotes the group of permutations of an L-element set and U_π is the unitary that permutes L Hilbert spaces according to the permutation π.

We now apply our recursive protocol, producing at each step an approximation of the desired M-copy state. Let us expand the states $|\psi\rangle^{\otimes N}$ and $|\psi\rangle^{\otimes M}$ as

$$|\psi\rangle^{\otimes L} = 2^{-L/2} \sum_{m=-L}^{L} \sqrt{\binom{L}{\frac{L-m}{2}}} |L, m\rangle \quad L = M, N,$$

then use the formulas for the fidelity and success probability derived in Sec. IV. At the first step of the protocol, the successful quantum operation produces an output state with the maximum possible fidelity, given by

$$F_{\max}^{(1)} = \frac{1}{2^M} \sum_{n=-N}^{N} \binom{M}{\frac{M-n}{2}}.$$
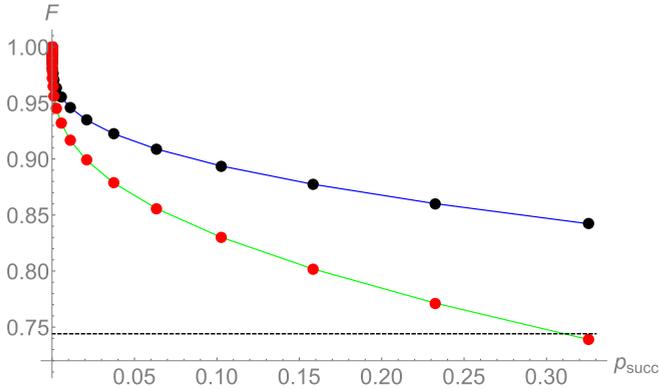
FIG. 6. Energy-preserving cloning of quantum clocks via the recursive protocol and its coherent coarse-graining. The figure shows the trade-off between success probability and fidelity for the N-to-M cloning of the clock state $|\psi_t\rangle = (e^{i\omega t/2}|0\rangle + e^{-i\omega t/2}|1\rangle)/\sqrt{2}$ with $M = 400$ and $N = 80$. The green solid line (with numerics represented by red dots) shows the probability-fidelity trade-off for a recursive protocol with $K = 32$ rounds. The blue solid line (with numerics represented by the black dots) shows the trade-off for filters generated by coherent coarse-graining, with the Tth point corresponding to the coherent coarse-graining of the first T steps of the recursive protocol. Finally, the black dashed line represents the fidelity for the optimal deterministic cloning protocol. Notice that the recursive protocol maintains fidelity larger than the optimal deterministic fidelity for all steps up to the last.

The above fidelity turns out to be equal to the absolute maximum of the fidelity achievable over all covariant quantum operations, derived by Fiurášek in Ref. [29]. For large N, the fidelity is close to 1 whenever M is small compared to N^2, thus allowing one to achieve quantum super-replication [27]. It is well known that the price of super-replication is a probability of success vanishing exponentially fast with N [27]. The main interest of our recursive protocol lies in the fact that it allows us to increase the probability of success. In a protocol with $K > 1$ steps, the average fidelity decreases at each step, while the probability of success increases. The trade-off between the fidelity and the probability of success is illustrated in Fig. 6 for the case of $N = 80$, $M = 400$, and $K = 32$, using Eqs. (55) and (56). In addition, we compare the fidelity of the recursive protocol with that of its coherent coarse-graining, given by Eq. (64). As already observed, the coherent coarse-graining achieves a higher fidelity, while keeping the same success probability. In the figure we also plot the optimal fidelity in the deterministic case (black dashed line in Fig. 6). The deterministic fidelity (derived in Theorem 3) coincides with the fidelity for phase-covariant cloning [66], meaning that the optimal cloner can be realized in an energy-preserving fashion.

Figure 6 well illustrates the advantages of the recursive protocol. At the first round the fidelity is very high, but the success probability has the extremely tiny value $p_{\text{succ}}^{(1)} = 6 \times 10^{-20}$. The subsequent rounds of the protocol increase the success probability dramatically, reaching a probability of approximately 23% at the 31st step. The fidelity for the recursive protocol remains higher than the optimal deterministic fidelity up to almost the very last step. An even better performance is attained through coherent coarse-graining.

D. Probabilistic energy-preserving amplification of coherent light

In quantum optics the energy-preserving instruments are those that preserve the average photon number. In the single-mode scenario, the number observable is nondegenerate and the energy-preserving quantum operations have diagonal Kraus operators in the Fock basis $\{|n\rangle\}$. In the following we consider the application of the recursive protocol to the amplification of the coherent state of light

$$|r_1\rangle \longrightarrow |r_2\rangle \quad 0 \leqslant r_1 \leqslant r_2.$$

Note that, since we require the amplification map to be part of a number-preserving quantum instrument, our protocol defines a phase-insensitive amplifier [84], which works equally well for the transition

$$|r_1 e^{i\theta}\rangle \longrightarrow |r_2 e^{i\theta}\rangle \quad 0 \leqslant r_1 \leqslant r_2,$$

where θ is an arbitrary angle.

Amplifying a coherent state without increasing its photon number seems to be a daunting task. However, the fact that the number is preserved only on average grants us the opportunity to reach high fidelity in a probabilistic fashion. In the case of amplifiers, the trade-off between success probability and fidelity is essentially a trade-off between success probability and photon number modulation.

Since the Hilbert space is infinite dimensional, our recursive protocol cannot be applied directly. To overcome the obstacle, we define a threshold N and assume that the successful operations project the input state inside the subspace spanned by Fock states with number smaller than N. Practically, for $N \gg r_2^2$, the projection can be done without affecting the fidelity. The fidelity at the kth round, given by Eq. (55), can be lower bounded as

$$F_{\text{max}}^{(k)} \geqslant 1 - e^{-r_2^2} \left(\frac{r_2^2 e}{N - k + 1} \right)^{N-k+1} \tag{78}$$

when $N - k + 1 > r_2^2$. On the other hand, the probability of success in Eq. (56) can be expressed as

$$p_{\text{succ}}^{(k)} = \begin{cases} e^{r_2^2 - r_1^2} \left(\frac{r_1}{r_2} \right)^{2N} F_{\text{max}}^{(1)} & k = 1, \\ e^{r_2^2 - r_1^2} \left(\frac{r_1}{r_2} \right)^{2N-2k+2} \left[1 - \left(\frac{r_1}{r_2} \right)^2 \right] F_{\text{max}}^{(k)} & k > 1. \end{cases} \tag{79}$$

Interestingly, the successful quantum operation at the first round of our protocol ($k = 1$) coincides with the optimal probabilistic amplifier for coherent states with known amplitude [29,85], which indeed can be implemented with energy-preserving operations. Specifically, evaluating Eqs. (78) and (79) for $k = 1$ one retrieves the expressions for the optimal fidelity and success probability appearing in Eqs. (6.24) and (6.36) of Ref. [85] For the subsequent rounds of the recursive protocol ($k > 1$), the input state is not coherent anymore and the successful quantum operation differs from the optimal coherent-state amplifier.

In Fig. 7 we show the performance of the recursive protocol and its coherent coarse-graining for the amplification of coherent states from $r_1 = 1$ to $r_2 = 1.5$. The threshold in the Fock space is chosen to be $N = 80$ and the protocol is applied recursively for $K = 81$ rounds. From the plot it
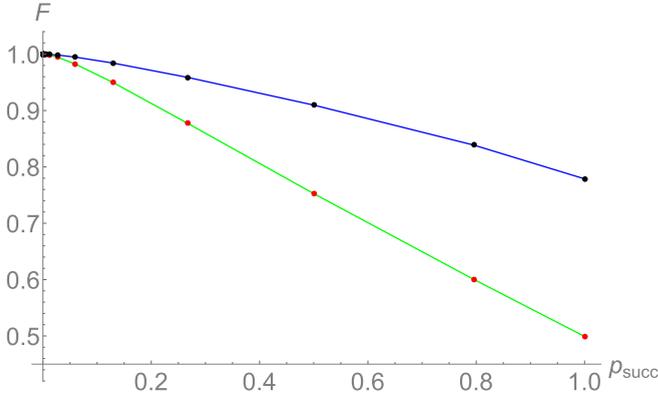
FIG. 7. Energy-preserving amplification of coherent light via the recursive protocol and its coherent coarse-graining. The figure shows the trade-off between success probability and the average fidelity for the amplification $|r_1\rangle \to |r_2\rangle$ with $r_1 = 1, r_2 = 1.5$, and $N = 80$. The green solid line (with numerics represented by red dots) shows the probability-fidelity trade-off for a recursive protocol with $K = 81$ rounds. The blue solid line (with numerics represented by the black dots) shows the trade-off for filters generated by coherent coarse-graining. Note that the difference between the two curves becomes large as the success probability tends to 1. For unit probability the recursive protocol and its coherent coarse-graining give fidelities $F_{\text{det}} = 49.9\%$ and $F'_{\text{det}} = |\langle r_1|r_2\rangle|^2 = 77.9\%$, respectively.

can be seen that the filters generated by coherent coarse-graining reach a relatively high fidelity, compatibly with the strong constraint set by the number-preserving condition. For instance, the coherent coarse-graining of the recursive protocol with $K = 80$ succeeds with probability $p_{\text{succ}} = 79.6\%$ and reaches fidelity $F = 83.9\%$. When the probability reaches 1, the fidelities of the recursive protocol and its coherent coarse-graining become $F_{\text{det}} = 49.9\%$ and $F'_{\text{det}} = |\langle r_1|r_2\rangle|^2 = 77.9\%$, respectively. The latter is well above the fidelity of the optimal amplifier for *arbitrary* coherent states, which is given by $F_{\text{universal}} = 4/9$ [31,86].

E. Energy-preserving correction in ancilla-driven quantum computation

In ancilla-driven quantum computation [75] the evolution of the system is determined by the outcomes of measurements on the ancilla. Ideally, the goal is to implement measurements that induce unitary gates on the system. To achieve this goal, the measurements should not extract any information about the state of the system: the probability of each outcome should be the probability that a particular unitary gate is applied to the system [87]. However, in many nonideal situations the measurement extracts some information, thus inducing a nonunitary evolution on the system. When this is the case, one can attempt to correct the unwanted nonunitarity by performing additional measurements. This type of correction has been studied in Refs. [88,89], where a number of different strategies have been proposed.

Here we consider the problem in the energy-preserving setting: suppose that a quantum system with d nondegenerate energy levels interacts with an ancilla via an energy-preserving unitary evolution. Then, the ancilla undergoes the measure-

ment of an observable that is compatible with the energy. As a result, the system evolves randomly according to an energy-preserving instrument $\{\mathcal{M}_x\}_{x \in \mathsf{X}}$. We assume that the measurement on the ancilla is a rank-1 projective measurement and, therefore, the quantum operations $\{\mathcal{M}_x\}$ are pure. For every given $x \in \mathsf{X}$, the problem is to correct the quantum operation $\mathcal{M}_x$, making it as close as possible to a desired energy-preserving unitary gate U_x. As a correction we allow ourselves to use an energy-preserving filter, with quantum operations $\{\mathcal{N}_{\text{succ}}^{(x)}, \mathcal{N}_{\text{fail}}^{(x)}\}$. Due to the presence of the filter, an initial pure state $|\eta\rangle$ is transformed probabilistically into the pure state

$$|\eta_x\rangle = \frac{N_{\text{succ}}^{(x)} M_x |\eta\rangle}{\left\| N_{\text{succ}}^{(x)} M_x |\eta\rangle \right\|}.$$

To evaluate the quality of the correction, we consider the fidelity between $|\eta_x\rangle$ and the target state $U_x|\eta\rangle$, averaging over all possible pure input states. Assuming that initially the state $|\eta\rangle$ is drawn at random according to the Haar measure, the conditional probability distribution over the pure states is given by

$$p(\eta|x, \text{succ})d\eta = \lambda_x \left\| N_{\text{succ}}^{(x)} M_x |\eta\rangle \right\|^2 d\eta,$$

where M_x and $N_{\text{succ}}^{(x)}$ are the Kraus operators of $\mathcal{M}_x$ and $\mathcal{N}_{\text{succ}}^{(x)}$, respectively, and λ_x is the normalization constant $\lambda_x := (\int \| N_{\text{succ}}^{(x)} M_x |\eta'\rangle \|^2 d\eta')^{-1}$. Hence, the average fidelity over all pure states is given by

$$
\begin{aligned}
F_x &:= \int d\eta\, p(\eta|x, \text{succ}) |\langle \eta_x | U_x | \eta \rangle|^2 \\
&= \frac{\int d\eta \left| \langle \eta | U_x^\dagger N_{\text{succ}}^{(x)} M_x |\eta\rangle \right|^2}{\int d\eta' \left\| N_{\text{succ}}^{(x)} M_x |\eta'\rangle \right\|^2} \\
&= \frac{F_0^{(x)} \cdot d + 1}{d + 1},
\end{aligned}
\tag{80}
$$

where $F_0^{(x)}$ is the fidelity given by

$$F_0^{(x)} = \frac{\left| \langle e_0 | U_x^\dagger N_{\text{succ}}^{(x)} M_x | e_0 \rangle \right|^2}{\left\| N_{\text{succ}}^{(x)} M_x | e_0 \rangle \right\|^2}$$

$$|e_0\rangle = \frac{\sum_{n=1}^{d} |n\rangle}{\sqrt{d}}.$$

Maximizing the average fidelity is then equivalent to finding the optimal quantum operation for the transformation

$$|\varphi_x\rangle := \frac{M_x |e_0\rangle}{\| M_x |e_0\rangle \|} \longrightarrow |\psi_x\rangle := U_x |e_0\rangle.$$

The maximization under the energy-preserving constraint is exactly the problem solved in this paper. In particular, for every outcome x we can use our recursive protocol to obtain a high-fidelity approximation of the desired transformation. In this context, it is immediate to realize that our protocol provides an approximate correction strategy, with the property that the overall quantum operation acts exactly like the target gate U_x in a suitable subspace, whose dimension shrinks at every step.
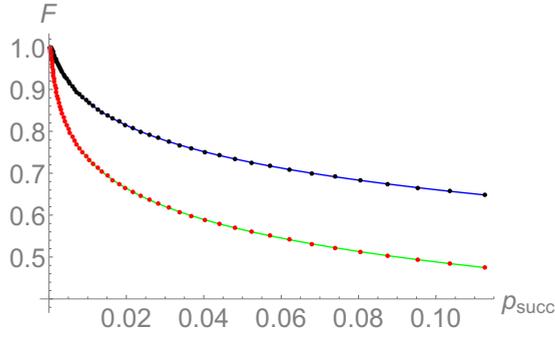
FIG. 8. Energy-preserving correction of a quantum operation via the recursive protocol and its coherent coarse-graining. The figure shows the trade-off between success probability and average fidelity for unlearning the quantum operation with Kraus operator $M_x = \sum_{n=1}^{d} \mu^{n/2}|n\rangle\langle n|$ with $d = 100$ and $\mu = 0.9$. The green solid line (with numerics represented by red dots) shows the probability-fidelity trade-off for a recursive protocol with $K = 70$ rounds. The blue solid line (with numerics represented by the black dots) shows the trade-off for filters generated by coherent coarse-graining.

For concreteness, let us see explicitly how the protocol works in a concrete example. We choose the quantum operation $\mathcal{M}_x$ with Kraus operator $M_x = \sum_{n=1}^{d} \mu^{n/2}|n\rangle\langle n|$. The fidelity at the kth step is given by

$$F_x^{(k)} = \frac{d + 2 - k}{d + 1},$$

while the probability of success, averaged over all pure states, is given by

$$p_{\text{succ}}^{(k)} = \begin{cases} \mu^{d-k}(1-\mu)^2(d+1-k)/(1-\mu^d) & k > 1, \\ \mu^{d-1}(1-\mu)d/(1-\mu^d) & k = 1. \end{cases}$$

The features of the recursive protocol and of its coherent coarse-graining are illustrated in Fig. 8 for $d = 100$, $\mu = 0.9$, and $K = 70$. The probability of success increases from a very small value ($p_{\text{succ}}^{(1)} = 3 \times 10^{-4}$) to approximately 14% at the 68th step, at the cost of a reduced fidelity.

VI. EXTENSION TO MIXED STATES

So far we considered transitions between pure states. However, for many practical applications it is important to consider transformations where the input states are mixed, e.g., due to the presence of decoherence. Normally the target state is still pure, since ideally one would like to remove the noise from the output. This is the case for tasks like mixed-state purification [32,40,41], super-broadcasting [42], and for the evaluation of the corresponding quantum benchmarks [90]. In all these cases, our techniques provide general upper bounds on the fidelity of the optimal energy-preserving operations. The bounds are achievable for a quite large class of states, which includes all the thermal states associated with stoquastic Hamiltonians [91,92], such as the Hamiltonians of flux qubits in Josephson junctions [93] and the Hamiltonians of (bosonic) Bose-Einstein condensates [94].

A. Deterministic transitions and eigenstate alignment

Given a generic mixed state ρ, we search for the best energy-preserving approximation of the transition $\rho \to |\psi\rangle\langle\psi|$, where $|\psi\rangle$ is a target pure state. To this purpose, it is convenient to decompose the initial state into blocks corresponding to the different energy eigenvalues, namely

$$\rho = \sum_{E,E'} \rho_{E,E'} \quad \rho_{E,E'} = P_E \rho P_{E'}. \tag{81}$$

With this notation, we have the following:

Theorem 4. For $p_{\text{succ}} = 1$, every energy-preserving approximation of the transition $\rho \to |\psi\rangle\langle\psi|$ satisfies the bound

$$F_{\text{det}} \leqslant \sum_{E,E' \in \mathsf{Sp}(\rho)} \sqrt{q_E q_{E'}} \|\rho_{E,E'}\|_1, \tag{82}$$

where $\|A\|_1 = \text{Tr}[\sqrt{A^\dagger A}]$ is the trace norm, $q_E = \|P_E|\psi\rangle\|^2$, and $\mathsf{Sp}(\rho)$ is the *energy spectrum of ρ*, defined as

$$\mathsf{Sp}(\rho) = \{E | P_E \rho P_E \neq 0\}. \tag{83}$$

The bound is achievable if the input state ρ is *block positive*, that is, if there exist orthonormal bases for the energy eigenspaces such that, when the matrix elements are taken in those bases, each matrix $[\rho_{E,E'}]$ is positive semidefinite. In this case, every quantum channel $\mathcal{A}$ satisfying the condition

$$\mathcal{A}(\rho) = \sum_{E,E'} \|\rho_{E,E'}\|_1 |\psi_E\rangle\langle\psi_{E'}| \quad \forall E, E' \in \mathsf{Sp}(\rho) \tag{84}$$

is optimal where $|\psi_E\rangle$ and $|\psi_{E'}\rangle$ are defined by Eq. (35).

Note that when the input state is pure, the bound (82) coincides with the optimal fidelity. We refer to every channel $\mathcal{A}$ satisfying Eq. (84) as an *eigenstate alignment of the mixed state ρ to the pure state $|\psi\rangle$*. Note that eigenstate alignment may not be a unitary operation anymore, because it may have to send different eigenstates with energy E to the fixed eigenstate $|\psi_E\rangle$.

The proof of Theorem 4 is provided in Appendix K. Three applications of the theorem are as follows:

Example 3 (nondegenerate Hamiltonians). When the Hamiltonian H is nondegenerate, the bound (82) becomes

$$F \leqslant \sum_{E,E'} \sqrt{q_E q_{E'}} |\langle \varphi_E | \rho | \varphi_{E'} \rangle|, \tag{85}$$

where $\{|\varphi_E\rangle\}$ is the energy eigenbasis. Note that since the Hamiltonian is nondegenerate, the choice of eigenbasis is unique up to phase transformations $|\varphi_E\rangle \mapsto |\varphi'_E\rangle = e^{i\theta_E}|\varphi_E\rangle$. The bound is achievable if, for a suitable choice of phases, one has

$$\langle \varphi_E | \rho_E | \varphi_{E'} \rangle \geqslant 0 \quad \forall E, E'. \tag{86}$$

Mixed states of this form were called *pure in phase* by D'Ariano *et al.* [95], who considered them in the context of phase estimation. Such states play an important role in the area of quantum Hamiltonian complexity, where they arise as thermal states of *stoquastic Hamiltonians* [91,92], i.e., Hamiltonians with nonpositive matrix elements in a given basis. Physically, we can consider a scenario where the system starts in the thermal state of a stoquastic Hamiltonian and subsequently undergoes a rapid change of Hamiltonian to

a diagonal one, making the initial thermal state a nontrivial resource.

Example 4 (pure states subject to random time evolution). Suppose that the system, initially prepared in a pure state $|\varphi\rangle$, has evolved under its free Hamiltonian for a time t which is not perfectly known, e.g., due to the finite-time resolution of the clocks available in the laboratory. Then, the system is effectively described by the mixed state

$$\rho = \int dt\, \pi(t) U_t |\varphi\rangle\langle\varphi| U_t^\dagger, \quad U_\tau = e^{-it H_{\text{sys}}/\hbar},$$

$\pi(t)$ being the probability distribution for t. In this case, the bound (82) becomes

$$F \leqslant \sum_{E,E'} \sqrt{q_E q_{E'} p_E p_{E'}} |\widetilde{\pi}(E - E')|,$$

where p_E and q_E are defined as in Eq. (34) and $\widetilde{\pi}$ is the Fourier transform of π. If the Fourier transform is positive (i.e., if the noise has positive spectrum), then the bound is attainable by every unitary operation U satisfying the eigenstate alignment condition [Eq. (36)].

Example 5 (multiple copies of qubit mixed states). Consider a system of N noninteracting qubits, each having the same Hamiltonian $H = E_0 Z/2$. Then, the total Hamiltonian of the system is degenerate and has the block diagonal form

$$H_{\text{sys}} = E_0 \bigoplus_l \left(J_z^{(l)} \otimes I_{d_l^{(N)}} \right),$$

where l is the quantum number of the total angular momentum, $J_z^{(l)}$ is the z component of the angular momentum operator in the subspace with quantum number l, and $I_{d_l^{(N)}}$ is the identity on a multiplicity space $\mathcal{M}_l^{(N)}$, of dimension

$$d_l^{(N)} = \frac{4l+2}{N+2l+2} \binom{N}{N/2+l}.$$

From the above decomposition it is clear that the eigenvalues of the energy are given by $E_m = E_0 m$, where m are the eigenvalues of the z component of the total angular momentum operator. A basis for the corresponding eigenspace is given by the vectors

$$|\varphi_{m,l,n}\rangle = |l,m\rangle \otimes |\mu_{l,n}\rangle,$$

where l goes from $|m|$ to $N/2$, $|l,m\rangle$ is the eigenstate of $J_z^{(l)}$ with eigenvalue m, and $\{|\mu_{l,n}\rangle | n = 1,\ldots,d_l^{(N)}\}$ is a basis for the multiplicity space. Now, suppose that each qubit is initially prepared in the state

$$\omega = \frac{e^{\beta X}}{\text{Tr}[e^{\beta X}]}, \quad \beta \geqslant 0, \quad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

With this choice, the state $\omega^{\otimes N}$ satisfies the condition for the achievability of the bound (82); indeed, one has

$$P_{E_m}\omega^{\otimes N}P_{E_{m'}} = \bigoplus_l \frac{\langle l,m|e^{2\beta J_x^{(l)}}|l,m'\rangle}{\text{Tr}[e^{\beta X}]^N} \left(|l,m\rangle\langle l,m'| \otimes I_{d_l^{(N)}} \right),$$

$$\tag{87}$$

and $\langle l,m|e^{2\beta J_x^{(l)}}|l,m'\rangle \geqslant 0$ since the matrix $J_x^{(l)}$ has positive matrix elements. Hence, the matrix elements of the operator

$P_{E_m}\rho P_{E_{m'}}$ in the basis $\{|l,m\rangle \otimes |\mu_{l,n}\rangle | l \geqslant \max\{|m|,|m'|\}, n = 1,\ldots d_l^{(N)}\}$ form a non-negative matrix. For the transition $\omega^{\otimes N} \rightarrow |\psi\rangle\langle\psi|$, eigenstate alignment is not a unitary operation, because all basis vectors with energy E_m are transformed into $|\psi_{E_m}\rangle$.

B. The ultimate probabilistic fidelity

We now provide the exact value of the maximum fidelity for the transition $\rho \rightarrow |\psi\rangle\langle\psi|$ when no restriction is imposed on the probability of success.

Theorem 5. For a finite-dimensional Hilbert space, the maximum fidelity over all energy-preserving operations is given by

$$F_{\max} = \|A\|_\infty,$$

$$A = \sum_{E,E'} \sqrt{q_E q_{E'}} |\psi_E\rangle\langle\psi_{E'}| \otimes \left(\rho_{E,E}^T\right)^{-\frac{1}{2}} \rho_{E,E'}^T \left(\rho_{E',E'}^T\right)^{-\frac{1}{2}},$$

$$\tag{88}$$

where $\|A\|_\infty = \max_{\||\psi\rangle\|=1} \|A|\psi\rangle\|$ denotes the operator norm and ρ^T denotes the transpose of ρ. For a quantum operation achieving fidelity $F_{\max}$, the maximum probability of success is equal to

$$p_{\text{succ}}^{\max} = \max_\sigma \min_E \frac{1}{\left\| \left(\rho_{E,E}^T\right)^{-1/2} \sigma_E \left(\rho_{E,E}^T\right)^{-1/2} \right\|_\infty}, \tag{89}$$

where the maximum $\max_\sigma$ runs over all density matrices σ with support contained in the eigenspace of A with maximum eigenvalue and $\sigma_E := \text{Tr}_1[(P_E \otimes I)\sigma]$, Tr_1 denoting the partial trace over the first Hilbert space.

The proof, provided in Appendix K, includes the explicit construction of the optimal strategy and an expression for the maximum probability of success.

C. Purification of coherence at zero energy cost

We now illustrate the application of our techniques to the concrete problem of purifying a mixed state [40,41]. Suppose that we are given N identical quantum systems, each prepared in the same mixed state, which happens to possess a nonzero amount of coherence across different energy levels. Can we collect the coherence present in the N systems and concentrate it in a single system? And can we do it without drawing energy from the outside? Mathematically, the task is to implement the transition

$$\omega^{\otimes N} \rightarrow |\phi\rangle\langle\phi| \otimes |\beta\rangle\langle\beta|,$$

where ω is the initial mixed state, $|\phi\rangle = \sum_{n=1}^d |n\rangle/\sqrt{d}$ is the maximally coherent state, and $|\beta\rangle$ is an eigenstate of the energy, in which $N-1$ systems are meant to be left. In the following we discuss the qubit case ($d = 2$) and we choose the mixed state to be

$$\omega = \frac{e^{\beta X}}{\text{Tr}[e^{\beta X}]}.$$

This state can be thought as the thermal state of the initial Hamiltonian $H_{\text{in}} = -X$ and represents a nontrivial resource if the Hamiltonian is suddenly changed into $H = Z$. For

(a)

(b)
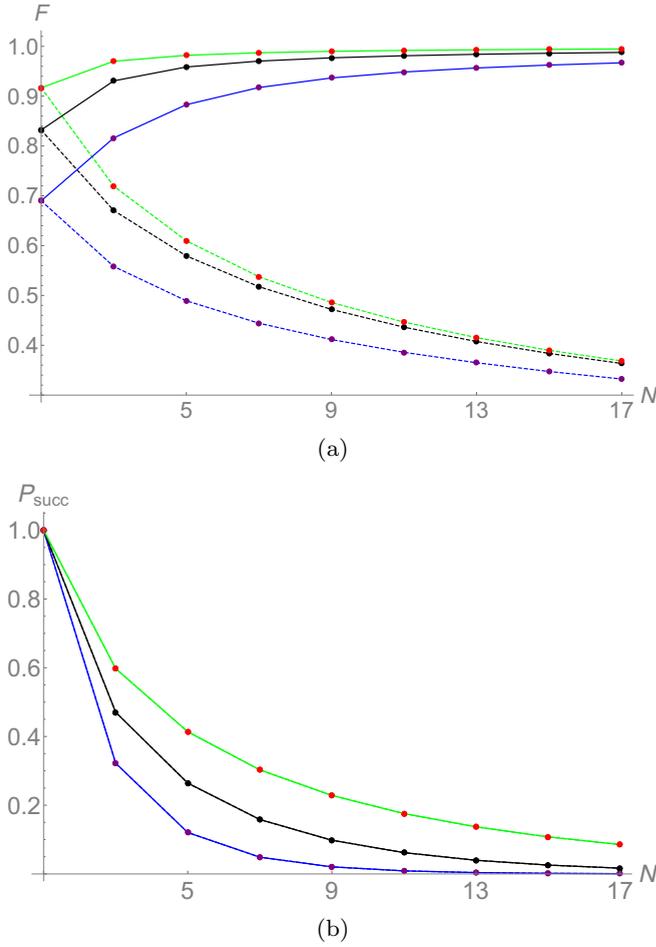
FIG. 9. Optimal purification of coherence via energy-preserving operations. Panel (a) shows the optimal fidelities as function of the number of input copies N. Here different colors represent different values of β—specifically, $\beta = 0.4$ (blue line with purple dots), $\beta = 0.8$ (black line with black dots), and $\beta = 1.2$ (green line with red dots). The dashed lines represent the deterministic fidelities, while the solid lines represent the probabilistic fidelities. While the probabilistic fidelities increase with N, the deterministic fidelities decrease, due to the restriction imposed by energy preservation. Panel (b) shows the maximum success probability for the quantum operations with maximum fidelity. The color code is the same as in panel (a).

simplicity, we choose N to be odd, so that $|\beta\rangle$ can be chosen to be an eigenstate with zero energy.

Let us consider first the deterministic transitions. The performance of the optimal energy-preserving channel is determined by Theorem 4, which leads to the expression

$$F_{\text{det}} = \frac{1}{2} \sum_{m,m'=\pm\frac{1}{2}} \sum_{l=\frac{1}{2}}^{\frac{N}{2}} \frac{d_l^{(N)} \langle l,m|e^{2\beta J_x^{(l)}}|l,m'\rangle}{\text{Tr}[e^{\beta X}]^N}.$$

Here the right-hand side follows from Eqs. (82) and (87), using the fact that $\omega^{\otimes N}$ is block positive, as observed in Example 5. The optimal deterministic fidelity is plotted in Fig. 9(a) for various values of N and β. Note that, quite counterintuitively, the deterministic fidelity *decreases* with the growth of N. The origin of this behavior can be found in the

constraint of energy preservation. Essentially, a deterministic energy-preserving operation cannot do anything better than realigning the blocks corresponding the values $m = \pm 1/2$ to the corresponding eigenstates. However, when N grows, the blocks are spread over an increasing number of values of m, so that the weight of the $m = \pm 1/2$ component becomes less and less significant. As a result, the deterministic fidelity vanishes in the limit $N \to \infty$. While the state $\omega^{\otimes N}$ contains an increasing amount of coherence, collecting this coherence from the high-energy blocks requires an exchange of energy with the surrounding environment.

For probabilistic strategies, the situation is different: the limitation due to energy-preservation can be partially lifted and the fidelity approaches unit as N increases. To evaluate the maximum fidelity, we have to compute the norm of the operator A defined in Eq. (88). In the case at hand, we have

$$A = \frac{1}{2} \sum_{m,m'=\pm\frac{1}{2}} \sum_{l=\frac{1}{2}}^{\frac{N}{2}} \frac{\langle l,m|e^{2\beta J_x^{(l)}}|l,m'\rangle}{\sqrt{\langle l,m|e^{2\beta J_x^{(l)}}|l,m\rangle \langle l,m'|e^{2\beta J_x^{(l)}}|l,m'\rangle}}$$

$$\times \left|\frac{1}{2},m\right\rangle\left\langle\frac{1}{2},m'\right| \otimes |l,m\rangle\langle l,m'| \otimes I_{d_l^{(N)}}.$$

Taking the operator norm, we then obtain

$$F_{\text{prob}} = \max_{l \in \{\frac{1}{2},...,\frac{N}{2}\}} \frac{1 + a_l}{2},$$

with

$$a_l = \frac{\langle l,\frac{1}{2}|e^{2\beta J_x^{(l)}}|l,-\frac{1}{2}\rangle}{\sqrt{\langle l,\frac{1}{2}|e^{2\beta J_x^{(l)}}|l,\frac{1}{2}\rangle\langle l,-\frac{1}{2}|e^{2\beta J_x^{(l)}}|l,-\frac{1}{2}\rangle}}.$$

The optimal probabilistic fidelity is plotted in Fig. 9(a) for different values of N and β.

Finally, Theorem 5 allows us to evaluate the maximum probability of success for the quantum operations that achieve maximum fidelity. According to the theorem, it is enough to characterize the density matrices that have support inside the eigenspace of A with maximum eigenvalue. Such density matrices have the form $\sigma = |\Phi_{N/2}\rangle\langle\Phi_{N/2}| \otimes \tau$, where $|\Phi_{N/2}\rangle$ is the maximally entangled state

$$|\Phi_{N/2}\rangle = \left(\left|\frac{1}{2},\frac{1}{2}\right\rangle\left|\frac{N}{2},\frac{1}{2}\right\rangle + \left|\frac{1}{2},-\frac{1}{2}\right\rangle\left|\frac{N}{2},-\frac{1}{2}\right\rangle\right)\Big/\sqrt{2}.$$

Hence, we have the relation

$$\sigma_m = \text{Tr}_1\left[\left(\left|\frac{1}{2},m\right\rangle\left\langle\frac{1}{2},m\right| \otimes I\right)\sigma\right]$$

$$= \frac{1}{2}\left|\frac{N}{2},m\right\rangle\left\langle\frac{N}{2},m\right| \otimes \tau,$$

which can be inserted into Eq. (89), yielding

$$p_{\text{succ}}^{\text{max}} = \max_\tau \min_{m=\pm\frac{1}{2}} \frac{2\langle l,m|e^{2\beta J_x^{(l)}}|l,m\rangle}{\text{Tr}[e^{\beta X}]^N \|\tau\|_\infty}$$

$$= \frac{2d_l^{(N)}\langle l,-\frac{1}{2}|e^{2\beta J_x^{(l)}}|l,-\frac{1}{2}\rangle}{\text{Tr}[e^{\beta X}]^N}.$$

A plot of the probability of success as a function of N and β is shown in Fig. 9(b).

VII. CONCLUSIONS

We introduced and analyzed the class of quantum operations that can be implemented with zero transfer of energy to the external environment. Within this class, we addressed the search for the optimal operations implementing a desired state transition. We considered operations that can generally be probabilistic, showing that the limitations arising from the preservation of the energy can be lifted to a surprising extent at the price of a reduced probability of success.

Our investigation started from the problem of transforming a given input state into a desired output state. To solve this problem, we developed two general techniques, dubbed eigenstate alignment and the Lüders reduction. Eigenstate alignment provides the best deterministic way to transform pure states at zero energy cost. The Lüders reduction applies more generally to the optimization of energy-preserving quantum operations. Essentially, it allows one to break down every quantum operation into the product of a pure probabilistic part followed by a deterministic part.

Employing these techniques, we reduced the search for the best energy-preserving transformations of pure states to a simple Lagrangian optimization.

The characterization of the optimal energy-preserving transformations of pure states allowed us to construct a multiple-round recursive protocol that achieves maximum fidelity with maximum success probability in each round. The probability of success of the protocol increases from one round to the next and, for a system with finite-energy spectrum, the protocol terminates deterministically in a finite number of steps.

Our protocol can be easily applied to every desired transformation of pure states, allowing one to find lower bounds to the optimal trade-off curve between fidelity and probability, whose exact expression is known only in a few cases. As an illustration of the versatility of the protocol, we applied it to a number of concrete tasks, including quantum phase estimation, cloning of coherence, energy-preserving amplification, and ancilla-driven computation at zero energy cost.

To further improve the bounds on the fidelity-probability trade-off, we applied the techniques of Lüders reduction and eigenstate alignment to the different histories resulting from subsequent rounds of the recursive protocol. Specifically, we introduced an operation of coherent coarse-graining, whereby a set of quantum operations are joined into a pure quantum operation, with the same probability of occurrence of the original set and, typically, with a higher fidelity with the target. Remarkably, when applied to the problem of phase estimation with maximally coherent states, coherent coarse-graining yields points that lie exactly on the optimal trade-off curve, provided that the number of energy levels is sufficiently large. Characterizing all the situations in which our method provides the optimal trade-off curve is an open problem. In general, we expect optimality to be achieved asymptotically in scenarios where the energy distribution of the state is sufficiently regular, including, e.g., the cases of phase estimation and quantum cloning in the asymptotic regime [96].

In this paper we provided a comprehensive study of the optimal quantum information processing under the energy-preserving constraint. A related avenue of future research is the study of optimal quantum information processing under general conservation laws. The techniques developed in this paper are already adapted to search of optimal quantum evolutions that preserve an *algebra* of quantum observables, such as the algebra generated by the angular momentum operators. Interactions that preserve the angular momentum have recently attracted attention in the implementation of quantum gates and quantum measurements [53,56,57,59,60], although the characterization of the optimal operations is still an open problem. In this context, our result suggests a strategy to approach the optimization, by considering probabilistic modulation of the amplitudes of the wave function in sectors with different angular momentum. Also in this case, our results allow one to construct first a recursive protocol and to increase its fidelity through the operation of coherent-coarse graining. While such generalizations are beyond the scope of the present paper, it is our hope that our work will pave the way to a systematic optimization of quantum operations under arbitrary conservation laws.

APPENDIX A: RELATION TO COVARIANT CHANNELS, HADAMARD CHANNELS, INCOHERENT CHANNELS, AND DECOHERENCE MAPS

Here we highlight the relations between energy-preserving channels and other important classes of channels considered in the literature. We start from the case of channels that are covariant under time evolution [63], i.e., channels that satisfy the relation

$$\mathcal{M}(U_t \cdot U_t^\dagger) = U_t \mathcal{M}(\cdot) U_t^\dagger \quad \forall t \in \mathbb{R}.$$

Covariant channels are at the basis of the resource theory of asymmetry [97–99].

Every covariant channel can be realized via a unitary evolution—as in Eq. (6)—with the property that the initial state of the environment $|\phi_0\rangle$ is an eigenstate of the energy and the joint unitary evolution preserves the sum of the energies of the system and the environment:

$$U^\dagger(H_{\text{sys}} + H_{\text{env}})U = H_{\text{sys}} + H_{\text{env}} \tag{A1}$$

(see, e.g., [100]). Note that the interaction preserves the sum of the energies of the system and the environment, but may involve an exchange of energy between them. This is the reason why not all covariant channels preserve the energy of the system individually.

A related subclass of covariant channels is the class of *cooling maps* considered in Ref. [101] which arise from the

exchange of energy between the system and an environment, originally initialized in the Gibbs state at near-zero temperature. In this case, the exchange of energy can go only in the direction from the system to the environment, as the environment can acquire energy from the system but not vice versa.

Energy-preserving channels are also closely related to Hadamard channels [102], also known as decoherence maps [103–105], and to the class of incoherent channels [33,34]. The relation can be seen in the case of a Hamiltonian H_{sys} with nondegenerate spectrum. In this case, the energy-preserving channels have been characterized in Ref. [104]; denoting by $\{|E_n\rangle\}_{n=1}^d$ the energy eigenbasis, it turns out that a channel is energy-preserving if and only if it is of the form

$$\mathcal{M}(\rho) = \sum_{m,n} C_{mn} \langle E_m|\rho|E_n\rangle |E_m\rangle\langle E_n|, \qquad \text{(A2)}$$

where $C = [C_{mn}]$ is a positive matrix with diagonal elements equal to 1. Channels of this are also known as *Hadamard channels* in the literature on quantum Shannon theory, where they represent one of the important classes of channels with tractable capacity regions [106,107]. Energy-preserving channels with nondegenerate Hamiltonian have a number of properties. First, note that every energy eigenstate is a fixed point of the channel, and so is every mixture of energy eigenstates. Hence, in the case of the nondegenerate Hamiltonian, the energy-preserving channels are a special case of *incoherent channels*, i.e., channels that transform incoherent mixtures into incoherent mixtures. Viewing coherence in the energy eigenbasis as a resource, it is clear that energy-preserving channels cannot be used to generate resourceful states from nonresourceful ones. On the contrary, typically they reduce quantum coherence, by damping down the off-diagonal elements of the density matrix [104].

The inclusion relations among energy-preserving, covariant, and incoherent operations are illustrated in Fig. 10.

APPENDIX B: PROOF OF THEOREMS 1 AND 2

Proof of Theorem 1. First, we show the implication $1 \implies 2$: every energy-preserving channel can be realized through a unitary interaction that separately preserves the energy of the system and the energy of the environment. We start from Eq. (14), namely $\mathcal{M}(\rho) = \text{Tr}_{\text{env}}[V\rho V^\dagger]$ with the isometry V given in Eq. (13). From the definition it follows that V can be written as

$$V = \bigoplus_E V_E \quad V_E := \sum_{k=1}^K P_E M_k P_E \otimes |\phi_k\rangle, \qquad \text{(B1)}$$

having chosen the environment to be K-dimensional, with $\mathcal{H}_{\text{env}} = \text{Span}\{|\phi_k\rangle | k = 1, \ldots, K\}$. Every operator V_E in Eq. (B1) is a unitary mapping from the eigenspace of H_{sys} with eigenvalue E, denoted by $\mathcal{H}_E$, to a subspace of $\mathcal{H}_E \otimes \mathcal{H}_{\text{env}}$, denoted by $\mathcal{S}_E$. Since $\mathcal{H}_{\text{env}}$ is K-dimensional, $\mathcal{H}_E \otimes \mathcal{H}_{\text{env}}$ can be decomposed as

$$\mathcal{H}_E \otimes \mathcal{H}_{\text{env}} = \bigoplus_{k=1}^K \mathcal{S}_E^{(k)}, \qquad \text{(B2)}$$
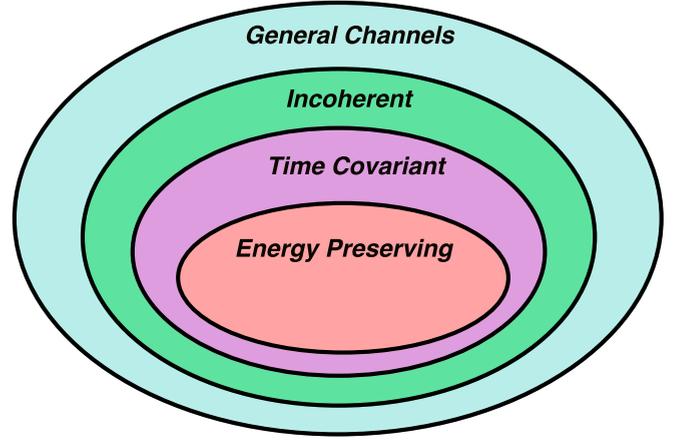
FIG. 10. Hierarchy of quantum channels. Energy-preserving channels are contained in the class of time-covariant channels, corresponding to the special case of channels that arise from an interaction that leaves the environment inside a fixed eigenspace at every time. For Hamiltonians with nondegenerate spectrum, covariant channels are a special case of incoherent channels, i.e., channels that do not generate coherence across the eigenbasis of the energy.

where $\{\mathcal{S}_E^{(k)}\}_{k=1}^K$ are orthogonal subspaces isomorphic to $\mathcal{S}_E$ and $\mathcal{S}_E^{(1)} \equiv \mathcal{S}_E$. Let us denote by $U_{E,k}$ the unitary that maps the subspace $\mathcal{S}_E$ into $\mathcal{S}_E^{(k)}$. With this notation, we can define the unitary operator U_E as

$$U_E|\psi\rangle|\phi_k\rangle := \begin{cases} V_E|\psi\rangle & k = 1, \\ U_{E,k}V_E|\psi\rangle & k > 1. \end{cases} \qquad \text{(B3)}$$

Note that, by construction, the operator U_E maps the subspace $\mathcal{H}_E \otimes \mathcal{H}_{\text{env}}$ into itself, and therefore satisfies the condition

$$U_E^\dagger P_E U_E = P_E. \qquad \text{(B4)}$$

Now, consider the unitary $U := \bigoplus_E U_E$. Clearly, U preserves the energy of the system; indeed, we have

$$U^\dagger H_{\text{sys}} U = U^\dagger \left(\bigoplus_E E P_E \right) U = \bigoplus_E E U_E^\dagger P_E U_E$$

$$= \bigoplus_E E P_E = H_{\text{sys}},$$

having used Eq. (B4) and the definition of U. On the other hand, choosing the Hamiltonian of the environment to be constant, e.g., $H_{\text{env}} = 0$, we trivially satisfy the condition $U^\dagger H_{\text{env}} U = H_{\text{env}}$. Finally, Eq. (B3) gives the relation

$$U|\psi\rangle|\phi_1\rangle = \bigoplus_E U_E|\psi\rangle|\phi_1\rangle$$

$$= \bigoplus_E V_E|\psi\rangle$$

$$= V|\psi\rangle$$

$$\forall|\psi\rangle \in \mathcal{H}_{\text{sys}}.$$

In turn, this implies the condition

$$\text{Tr}_{\text{env}}[U(|\psi\rangle\langle\psi| \otimes |\phi_1\rangle\langle\phi_1|)U^\dagger]$$

$$= \text{Tr}_{\text{env}}[V|\psi\rangle\langle\psi|V^\dagger] = \mathcal{M}(|\psi\rangle\langle\psi|),$$

meaning that the unitary U implements the desired channel $\mathcal{M}$ on every pure state, and, by linearity, on every mixed state.

Let us now show the implication $2 \implies 3$: the unitary evolution U can be realized through an interaction $H_{\mathrm{int}}(t)$ that commutes with the Hamiltonians H_{sys} and H_{env} at all times. The proof is immediate from the spectral decomposition

$$ U = \sum_n e^{-i\theta_n} |\phi_n\rangle\langle\phi_n|, \quad \theta_n \in [0, 2\pi). $$

Indeed, it is enough to define

$$ H := \frac{i\hbar}{t_1 - t_0} \ln U, \quad \ln U = \sum_n -i\theta |\varphi_n\rangle\langle\varphi_n|, $$

and

$$ H_{\mathrm{int}}(t) := g(t) H, $$

where $g(t)$ is an arbitrary function that quantifies the strength of the interaction and satisfies the conditions

$$ g(t) = 0 \quad \forall t \leqslant t_0, \quad \forall t \geqslant t_1, $$
$$ g(t) \geqslant 0 \quad \forall t \in (t_0, t_1), $$
$$ \int_{t_0}^{t_1} dt\, g(t) = 1. $$

Finally, we note that the implication $3 \implies 1$ has been already proven in the main text. ∎

Note that the construction in the above proof allows one to engineer a *minimal* realization of the desired channel, in which the Hamiltonian of the environment is fully degenerate. However, nothing prevents us from regarding the Hilbert space of the environment in the above construction as a degenerate eigenspace of a nontrivial Hamiltonian, acting on a larger space.

Proof of Theorem 2. For every x, let us pick a Kraus decomposition $\mathcal{M}_x(\rho) = \sum_{k=1}^{m_x} M_{x,k} \rho M_{x,k}^\dagger$. By definition, the operators $\{M_{x,k} | x \in \mathsf{X}, k = 1, \ldots, m_x\}$ are Kraus operators for the energy-preserving channel $\mathcal{M} = \sum_x \mathcal{M}_x$. Hence, the construction in the previous proof yields an environment with Hilbert space

$$ \mathcal{H}_{\mathrm{env}} = \mathsf{Span}\{|\phi_{x,k}\rangle | x \in \mathsf{X}, k = 1, \ldots, m_x\}, $$

a constant Hamiltonian H_{env}, a unit vector $|\phi_{1,1}\rangle \in \mathcal{H}_{\mathrm{env}}$, and a unitary operator U such that

$$ U|\psi\rangle|\phi_{1,1}\rangle = \sum_{x,k} M_{x,k}|\psi\rangle|\phi_{x,k}\rangle \quad \forall|\psi\rangle \in \mathcal{H}_{\mathrm{sys}} \quad \text{(B5)} $$

and U satisfies the conditions $U^\dagger H_{\mathrm{sys}} U = H_{\mathrm{sys}}$ and $U^\dagger H_{\mathrm{env}} U = H_{\mathrm{env}}$. Now, consider the projective measurement $\{Q_x\}_{x \in \mathsf{X}}$ defined by $Q_x = \sum_{k=1}^{m_x} |\phi_{x,k}\rangle\langle\phi_{x,k}|$. By Eq. (B5), one has the desired condition

$$ \mathrm{Tr}_{\mathrm{env}}[(I_{\mathrm{sys}} \otimes Q_x) U(|\psi\rangle\langle\psi| \otimes |\phi_{1,1}\rangle\langle\phi_{1,1}|) U^\dagger] $$
$$ = \mathcal{M}_x(|\psi\rangle\langle\psi|) $$

valid for every x and for every pure state $|\psi\rangle$, and, by linearity, for every mixed state ρ. The projective measurement $\{Q_x\}_{x \in \mathsf{X}}$ can be regarded as a measurement of an observable $O = \sum_{x \in \mathsf{X}} f(x) Q_x$, where $f : \mathsf{X} \to \mathbb{R}$ is a fixed (but otherwise arbitrary) injective function. Since the Hamiltonian of the environment is constant, the observable O trivially satisfies Yanase's condition (18). ∎

APPENDIX C: WEAK VS STRONG ENERGY PRESERVATION FOR PROBABILISTIC TRANSFORMATIONS

Here we compare our notion of energy-preserving transformation—defined as transformations that can be part of an energy-preserving instrument—with a stronger notion. For this reason, here we refer to our condition as *weak energy preservation*:

Definition 4. A probabilistic transformation $\mathcal{M}_0$ satisfies *weak energy preservation* if there exists an energy-preserving instrument $\{\mathcal{M}_x\}_{x \in \mathsf{X}}$ and an outcome $x_0 \in \mathsf{X}$ such that $\mathcal{M}_0 = \mathcal{M}_{x_0}$.

The above notion of weak energy preservation—on average over all the possible outcomes of a measurement—is fundamentally different from an alternative, strong notion of energy preservation, which requires the output of the transformation to have the same energy distribution of the input:

Definition 5. A probabilistic transformation $\mathcal{M}_0$ satisfies *strong energy preservation* if, for every input state ρ, one has

$$ \mathrm{Tr}\left[\rho' H_{\mathrm{sys}}^n\right] = \mathrm{Tr}\left[\rho H_{\mathrm{sys}}^n\right] \quad \forall n \in \mathbb{N}, \quad \text{(C1)} $$

where ρ' is the conditional output state $\rho' = \mathcal{M}_0(\rho)/\mathrm{Tr}[\mathcal{M}_0(\rho)]$.

While strong energy preservation may seem more natural at first sight, a closer inspection reveals that the condition is quite restrictive. For instance, it even forbids ideal energy measurements on the system, which are the prototype of measurements that can be performed at no energy cost [55,56,61]. Indeed, a von Neumann measurement of the energy trivially collapses a mixture of different energy eigenstates into a single energy eigenstate, thus violating Eq. (C1). Furthermore, it is not hard to see that for nondegenerate Hamiltonians, every quantum operation satisfying strong energy preservation must be proportional to a *deterministic* energy-preserving transformation. In other words, the operation can be realized by tossing a biased coin, and, if the coin turns out heads, by applying a deterministic transformation. As a result, the strong notion of energy-preservation would make the study of probabilistic transformations irrelevant.

In this work we focus on the more flexible notion of weak energy preservation, which captures exactly the class of transformations that can be implemented at no energy cost. Since there is no ambiguity, we will omit the specification "weak" and refer to these transformations simply as energy-preserving.

APPENDIX D: DECOMPOSITION OF STATIONARY COVARIANT INSTRUMENTS

Proof of Proposition 1. In Ref. [27] it was proven that every instrument $\{\mathcal{M}_x\}_{x \in \mathsf{X}}$ can decomposed as $\mathcal{M}_x = \mathcal{C}_x \mathcal{P}_x$, where $\{\mathcal{P}_x\}_{x \in \mathsf{X}}$ is the quantum instrument defined by

$$ \mathcal{P}_x(\cdot) = \sqrt{P_x} \cdot \sqrt{P_x}, \quad P_x = \mathcal{N}_x^\dagger(I) \quad \text{(D1)} $$

and $\mathcal{C}_x$ is a suitable quantum channel. Now, suppose that $\mathcal{M}_x$ is stationary and covariant, with

$$V_t \mathcal{M}_x(\cdot) V_t^\dagger = \mathcal{M}_x(U_t \cdot U_t^\dagger) \quad \forall t \in \mathbb{R}.$$

Then, one has

$$
\begin{aligned}
U_t^\dagger P_x U_t &= U_t^\dagger \mathcal{M}_x^\dagger(I) U_t \\
&= \mathcal{M}_x^\dagger(V_t^\dagger V_t) \\
&= \mathcal{M}_x^\dagger(I) \\
&= P_x \\
&\forall t \in \mathbb{R}.
\end{aligned}
$$

Taking derivatives with respect to t on both sides and recalling the definition, one then obtains

$$[P_x, H_{\text{sys}}] = 0 \tag{D2}$$

and consequently $[\sqrt{P_x}, H_{\text{sys}}] = 0$. Hence, the quantum operation

$$\mathcal{M}_x(\cdot) = \sqrt{P_x} \cdot \sqrt{P_x}$$

is energy-preserving (Lemma 1). It remains to show that the channel $\mathcal{C}_x$ is covariant. This can be done easily when P_x is invertible. In this case, one has $\mathcal{C}_x = \mathcal{M}_x \mathcal{P}_x^{-1}$, where $\mathcal{P}_x^{-1}$ is the completely positive map

$$\mathcal{P}_x^{-1}(\cdot) = \sqrt{P_x^{-1}} \cdot \sqrt{P_x^{-1}}.$$

Using this fact, we obtain

$$
\begin{aligned}
\mathcal{C}_x(U_t \cdot U_t^\dagger) &= \mathcal{M}_x \mathcal{P}_x^{-1}(U_t \cdot U_t^\dagger) \\
&= \mathcal{M}_x\left(\sqrt{P_x^{-1}} U_t \cdot U_t^\dagger \sqrt{P_x^{-1}}\right) \\
&= \mathcal{M}_x\left(U_t \sqrt{P_x^{-1}} \cdot \sqrt{P_x^{-1}} U_t^\dagger\right) \\
&= V_t \mathcal{M}_x\left(\sqrt{P_x^{-1}} \cdot \sqrt{P_x^{-1}}\right) V_t^\dagger \\
&= V_t \mathcal{C}_x(\cdot) V_t^\dagger \quad \forall t \in \mathbb{R};
\end{aligned}
$$

that is, $\mathcal{C}_x$ is covariant. When P_x is not invertible, the above reasoning shows that the restriction of $\mathcal{C}_x$ to the support of P_x is covariant. On the orthogonal complement of the support of P_x, the action of the channel $\mathcal{C}_x$ can be redefined arbitrarily in such a way that the covariance condition is guaranteed. ∎

APPENDIX E: DECOMPOSITION OF NONSTATIONARY COVARIANT INSTRUMENTS

Proof of Proposition 2. The proof follows the same lines as the proof of Proposition 1: Define the instrument $\{\mathcal{N}_y\}$ as

$$\mathcal{N}_y := \sum_{x \in \mathsf{X}} \mathcal{M}_x$$

(it is understood that, when the orbit is continuous, the sum has to be replaced by an integral over the orbit). Define the operators $P_y := \mathcal{N}_y^\dagger(I)$ and the pure instrument $\{\mathcal{P}_y\}$, with

$$\mathcal{P}_y(\cdot) = \sqrt{P_y} \cdot \sqrt{P_y}.$$

Then, the same argument used in the proof of Proposition 1 shows that the instrument $\{\mathcal{P}_y\}$ is energy-preserving. Then, for every $y \in \mathsf{Y}$, define the instrument $\{\mathcal{M}_x^{(y)}\}_{x \in \mathsf{O}_y}$ as

$$\mathcal{M}_x^{(y)} := \mathcal{M}_x\left(P_y^{-1/2} \cdot P_y^{-1/2}\right).$$

By definition, the quantum operation $\mathcal{M}_x^{(y)}$ satisfies the condition

$$\mathcal{M}_x = \mathcal{M}_x^{(y)} \circ \mathcal{P}_y. \tag{E1}$$

Moreover, the quantum operations $\{\mathcal{M}_x^{(y)}\}_{x \in \mathsf{O}_y}$ form a well-normalized instrument transforming states with support of P_y into states with support $\mathcal{H}_{\text{out}}$. When the support of P_y is not the whole of $\mathcal{H}_{\text{in}}$, one can complete the quantum operations $\{\mathcal{M}_x^{(y)}\}_{x \in \mathsf{O}_y}$ by extending their action on the kernel of P_y. As a result, one obtains a well-normalized instrument from $\mathcal{H}_{\text{in}}$ to $\mathcal{H}_{\text{out}}$, without affecting the validity of Eq. (E1). ∎

APPENDIX F: OPTIMALITY OF EIGENSTATE ALIGNMENT

Proof of Theorem 3. Let $\mathcal{M}$ be an energy-preserving quantum operation such that takes place deterministically on $|\varphi\rangle$. The, the fidelity of $\mathcal{M}$ for the transition $|\varphi\rangle \to |\psi\rangle$ can be computed as

$$
\begin{aligned}
F &= \langle\psi|\mathcal{M}(|\varphi\rangle\langle\varphi|)|\psi\rangle \\
&= \sum_{E,E'} \sqrt{p_E q_E p'_E q'_E} \langle\psi_E|\mathcal{M}(|\varphi_E\rangle\langle\varphi_{E'}|)|\psi_{E'}\rangle \\
&= \sum_{E,E'} \sqrt{p_E q_E p'_E q'_E} \langle\psi_E|\mathcal{M}(|\varphi_E\rangle\langle\varphi_{E'}|)|\psi_{E'}\rangle.
\end{aligned}
$$

Let $\mathcal{M}$ be an energy-preserving quantum operation, with Kraus decomposition $\mathcal{M}(\rho) = \sum_k M_k \rho M_k^\dagger$. Since $\mathcal{M}$ is energy-preserving, every Kraus operator M_k must be of the form

$$M_k = \bigoplus_E P_E M_k P_E; \tag{F1}$$

cf. Eq. (13). The probability of success is then given by

$$
\begin{aligned}
p_{\text{succ}} &= \text{Tr}[\mathcal{M}(|\varphi\rangle\langle\varphi|)] \\
&= \sum_k \langle\varphi|M_k^\dagger M_k|\varphi\rangle \\
&= \sum_k \sum_{E,E'} \sqrt{p_E p'_E} \langle\varphi_E|M_k^\dagger M_k|\varphi_{E'}\rangle \\
&= \sum_k \sum_E p_E \langle\varphi_E|M_k^\dagger M_k|\varphi_E\rangle, \tag{F2}
\end{aligned}
$$

the third equality following from the decomposition of $|\varphi\rangle$ [Eq. (34)] and the fourth equality following from the block diagonal form of M_k. Now, note that one has

$$\sum_k \langle\varphi_E|M_k^\dagger M_k|\varphi_E\rangle = \text{Tr}[\mathcal{M}(|\varphi_E\rangle\langle\varphi_E|)] \leqslant 1.$$

Hence, the condition $p_{\mathrm{succ}} = 1$ can be satisfied by Eq. (F2) only if the equality

$$\sum_k \langle \varphi_E | M_k^\dagger M_k | \varphi_E \rangle = 1 \tag{F3}$$

holds for every E such that $p_E \neq 0$. We will now optimize the fidelity subject to this constraint.

The fidelity for the transformation $|\varphi\rangle \to |\psi\rangle$ can be expressed as

$$
\begin{aligned}
F &= \langle \psi | \mathcal{C}(|\varphi\rangle\langle\varphi|) | \psi \rangle \\
&= \sum_k |\langle \psi | M_k | \varphi \rangle|^2 \\
&= \sum_k \left| \sum_E \langle \psi | P_E M_k P_E | \varphi \rangle \right|^2 \\
&= \sum_k \left| \sum_E \sqrt{p_E q_E} \langle \psi_E | M_k | \varphi_E \rangle \right|^2,
\end{aligned}
$$

having used the decompositions of Eq. (34). Then, one has the bound

$$
\begin{aligned}
F &\leqslant \sum_k \left(\sum_E \sqrt{p_E q_E} |\langle \psi_E | M_k | \varphi_E \rangle| \right)^2 \\
&\leqslant \sum_k \left(\sum_E \sqrt{p_E q_E} \| M_k | \varphi_E \rangle \| \right)^2 \\
&\leqslant \sum_k \left(\sum_E \sqrt{p_E q_E} \| M_k | \varphi_E \rangle \|^2 \right) \left(\sum_{E'} \sqrt{p_{E'} q_{E'}} \right),
\end{aligned}
$$

the third bound coming from the Schwarz inequality. Inserting the relation $\| M_k | \varphi_E \rangle \|^2 = \langle \varphi_E | M_k^\dagger M_k | \varphi_E \rangle$ one finally obtains

$$
\begin{aligned}
F &\leqslant \sum_k \left(\sum_E \sqrt{p_E q_E} \langle \varphi_E | M_k^\dagger M_k | \varphi_E \rangle \right) \left(\sum_{E'} \sqrt{p_{E'} q_{E'}} \right) \\
&= \left(\sum_E \sqrt{p_E q_E} \right)^2,
\end{aligned}
$$

having used Eq. (F3). By direct inspection it is immediate to see that eigenstate alignment attains the bound. ∎

APPENDIX G: DECOMPOSITION OF A QUANTUM OPERATION IN TERMS OF ITS LÜDERS REDUCTION

Proof of Proposition 3. In the case when P is a projector, the decomposition $\mathcal{M} = \mathcal{C} \circ \mathcal{P}$ was proven by Bartlett *et al.* in Ref. [108]. In general, the decomposition follows from the decomposition of an instrument proven in Ref. [27]. Now, suppose that $\mathcal{M}$ is energy-preserving. In this case, we have already seen that P commutes with H_{sys} and, therefore, the quantum operation $\mathcal{P}(\cdot) = \sqrt{P} \cdot \sqrt{P}$ is energy-preserving (see the proof of Proposition 1 in Appendix D). It only remains to prove that $\mathcal{C}$ is energy-preserving. This is easily done when P is invertible; in this case, one has

$$\mathcal{C} = \mathcal{M}\mathcal{P}^{-1}, \quad \mathcal{P}^{-1}(\cdot) = \sqrt{P^{-1}} \cdot \sqrt{P^{-1}}, \tag{G1}$$

meaning that every Kraus operator of $\mathcal{C}$ is the form $C_k = M_k \sqrt{P^{-1}}$, where M_k is a Kraus operator for $\mathcal{M}$. Now, since $\mathcal{M}$ is energy-preserving, every operator M_k commutes with H_{sys}. Hence, also C_k commutes with H_{sys}. Since C_k is a generic Kraus operator, Lemma 1 implies that $\mathcal{C}$ is energy-preserving. When P_x is not invertible, the above reasoning shows that the restriction of $\mathcal{C}$ to the support of P is energy-preserving. On the orthogonal complement of the support of P, the action of the channel $\mathcal{C}$ can be redefined arbitrarily in such a way that energy-preservation is guaranteed. ∎

APPENDIX H: THE ULTIMATE LIMITS OF PROBABILISTIC OPERATIONS

Proof of Proposition 5. By Proposition 4, the maximum fidelity is achieved by a pure quantum operation $\mathcal{M}(\cdot) = M \cdot M^\dagger$, with success probability

$$p_{\mathrm{succ}} = \sum_E p_E \langle \varphi_E | P_{\mathrm{succ}} | \varphi_E \rangle, \quad P_{\mathrm{succ}} = M^\dagger M, \tag{H1}$$

and fidelity

$$F = \left(\sum_E \sqrt{p'_E q_E} \right)^2, \quad p'_E = \frac{p_E \langle \varphi_E | P_{\mathrm{succ}} | \varphi_E \rangle}{p_{\mathrm{succ}}}. \tag{H2}$$

By the Schwarz inequality, one immediately obtains the bound

$$F \leqslant \sum_{E \in \mathsf{Sp}(\varphi) \cap \mathsf{Sp}(\psi)} q_E.$$

The bound is achieved if and only if

$$\langle \varphi_E | P_{\mathrm{succ}} | \varphi_E \rangle = c \frac{q_E}{p_E}, \quad \forall E \in \mathsf{Sp}(\varphi) \cap \mathsf{Sp}(\psi), \tag{H3}$$

for some constant $c \geqslant 0$. Note that since the $\langle \varphi_E | P_{\mathrm{succ}} | \varphi_E \rangle$ is the probability that $\mathcal{M}$ takes place on the state $|\varphi_E\rangle$, the constant c must satisfy the relation

$$c \leqslant \frac{p_E}{q_E} \quad \forall E \in \mathsf{Sp}(\varphi) \cap \mathsf{Sp}(\psi). \tag{H4}$$

Now, recall that the quantum operation $\mathcal{M}$ was constructed from the Lüders reduction and eigenstate alignment. Hence, its Kraus operator M must satisfy the condition

$$M |\varphi_E\rangle \propto |\psi_E\rangle \quad \forall E \in \mathsf{Sp}(\varphi) \cap \mathsf{Sp}(\psi). \tag{H5}$$

Combining this fact with Eq. (H3) and recalling that $P_{\mathrm{succ}} = M^\dagger M$, we obtain

$$M |\varphi_E\rangle = \sqrt{c \frac{q_E}{p_E}} |\psi_E\rangle \quad \forall E \in \mathsf{Sp}(\varphi) \cap \mathsf{Sp}(\psi). \tag{H6}$$

Now, inserting Eq. (H3) into Eq. (H1), the probability of success becomes

$$p_{\mathrm{succ}} = c \sum_{E \in \mathsf{Sp}(\varphi) \cap \mathsf{Sp}(\psi)} q_E. \tag{H7}$$

Given the constraint (H4), the maximum success probability is obtained by choosing

$$c = \min_{E \in \mathsf{Sp}(\varphi) \cap \mathsf{Sp}(\psi)} \frac{p_E}{q_E}. \tag{H8}$$

Inserting this value into Eqs. (H7) and (H6), we finally obtain the desired relations (49) and (50). ∎

APPENDIX I: DERIVATION OF THE OPTIMAL RECURSIVE PROTOCOL

The optimal filter in the kth round is determined by induction from the requirements 1–3 in Sec. IV A. At the first round, the filter attempts at converting $|\varphi\rangle$ into $|\psi\rangle$. According to Proposition 5, the maximum fidelity is given by

$$F_{\max}^{(1)} = \sum_{E \in \mathsf{Sp}(\varphi)} q_E \tag{I1}$$

and can be achieved with probability

$$p_{\mathrm{succ}}^{(1)} = \left(\min_{E \in \mathsf{Sp}(\varphi) \cap \mathsf{Sp}(\psi)} \frac{p_E}{q_E} \right) F_{\max}^{(1)}.$$

The optimal quantum operation must be pure and its Kraus operator $B_{\mathrm{succ}}^{(1)}$ must satisfy the condition

$$B_{\mathrm{succ}}^{(1)} |\varphi_E\rangle = |\psi_E\rangle \quad \forall E \in \mathsf{R}_1,$$

where R_1 is the set of energy values in $\mathsf{Sp}(\varphi) \cap \mathsf{Sp}(\psi)$ that minimize the ratio p_E/q_E [cf. Eq. (H6)]. Writing the unsuccessful quantum operation in the Kraus form $\mathcal{B}_{\mathrm{fail}}^{(1)}(\cdot) = \sum_t B_{\mathrm{fail},t}^{(1)} \cdot B_{\mathrm{fail},t}^{(1)\dagger}$, we then obtain

$$B_{\mathrm{fail},t}^{(1)} |\varphi_E\rangle = 0 \quad \forall E \in \mathsf{R}_1 \tag{I2}$$

for every possible value of t. At the second step, the filter attempts to produce the target state $|\psi\rangle$ from the state

$$\rho^{(2)} = \sum_t p_t^{(2)} |\varphi^{(2,t)}\rangle\langle\varphi^{(2,t)}|,$$

with $p_t^{(2)} := \|B_{\mathrm{fail},t}^{(1)} |\varphi\rangle\|^2 / \sum_{t'} \|B_{\mathrm{fail},t'}^{(1)} |\varphi\rangle\|^2$ and

$$|\varphi^{(2,t)}\rangle := \frac{B_{\mathrm{fail},t}^{(1)} |\varphi\rangle}{\|B_{\mathrm{fail},t}^{(1)} |\varphi\rangle\|}. \tag{I3}$$

Clearly, the maximum fidelity achievable from the state $\rho^{(2)}$ cannot be larger than the maximum over t of the fidelity achievable from $|\varphi^{(2,t)}\rangle$. Now, let us expand each state as

$$|\varphi^{(2,t)}\rangle = \sum_E \sqrt{p_E^{(2,t)}} |\varphi_E^{(2,t)}\rangle,$$

for suitable probabilities $\{p_E^{(2,t)}\}$ and suitable energy eigenstates $\{|\varphi_E^{(2,t)}\rangle\}$. Note that due to the condition in Eq. (I2), one has

$$\mathsf{Sp}(\varphi^{(2,k)}) \subseteq \mathsf{Sp}(\varphi) \setminus \mathsf{R}_1. \tag{I4}$$

Using this fact, we can upper-bound the fidelity achievable from the state $|\varphi^{(2,t)}\rangle$—call it $F_{\max}^{(2,t)}$—as

$$\begin{aligned} F_{\max}^{(2,t)} &\leqslant \sum_{E \in \mathsf{Sp}(\varphi^{(2,t)})} q_E \\ &\leqslant \sum_{E \in \mathsf{Sp}(\varphi) \setminus \mathsf{R}_1} q_E, \end{aligned}$$

the first inequality coming from Proposition 5. In turn, this allows us to upper-bound the overall fidelity at the second step as

$$F^{(2)} \leqslant \max_t F_{\max}^{(2,t)} \leqslant \sum_{E \in \mathsf{Sp}(\varphi) \setminus \mathsf{R}_1} q_E =: F_{\max}^{(2)}. \tag{I5}$$

The bound is attained when the quantum operation $\mathcal{B}_{\mathrm{fail}}^{(1)}$ is pure and its Kraus operator is given by

$$B_{\mathrm{fail}}^{(1)} := \sqrt{I - B_{\mathrm{succ}}^{(1)\dagger} B_{\mathrm{succ}}^{(1)}}. \tag{I6}$$

Luckily, this choice maximizes not only the fidelity at the second step but also the *probability* that maximum fidelity is achieved; indeed, denoting by $p_{\mathrm{succ}}^{(2)}$ the probability that the output has fidelity $F_{\max}^{(2)}$ with the target and by $p_{\mathrm{succ}}^{(2,t)}$ the probability that the optimal transformation takes place on the state $|\varphi^{(2,t)}\rangle$, we have the bound

$$\begin{aligned} p_{\mathrm{succ}}^{(2)} &\leqslant \sum_t p_t^{(2)} p_{\mathrm{succ}}^{(2,t)} \\ &= \sum_t p_t^{(2)} \left[\min_{E \in \mathsf{Sp}(\varphi^{(2,t)}) \cap \mathsf{Sp}(\psi)} \frac{p_E^{(2,t)}}{q_E} \right] F_{\max}^{(2,t)} \\ &\leqslant \min_E \left[\sum_t \frac{p_t^{(2)} p_E^{(2,t)}}{q_E} \right] F_{\max}^{(2)}, \end{aligned} \tag{I7}$$

the equality in the second line coming from Proposition 5. It is easy to verify that the pure quantum operation of Eq. (I6) reaches the bound; indeed, its output state

$$|\varphi^{(2)}\rangle = \frac{B_{\mathrm{fail}}^{(1)} |\varphi\rangle}{\|B_{\mathrm{fail}}^{(1)} |\varphi\rangle\|}$$

can be converted optimally into the state $|\psi\rangle$ with probability given by Proposition 5, which now yields

$$p_{\mathrm{succ}}^{(2)} = \min_{E \in \mathsf{Sp}(\varphi^{(2)}) \cap \mathsf{Sp}(\psi)} \left[\frac{p_E^{(2)}}{q_E} \right] F_{\max}^{(2)} \tag{I8}$$

with $p_E^{(2)} := \|P_E |\varphi^{(2)}\rangle\|^2 \equiv \sum_t p_t^{(2)} p_E^{(2,t)}$. Inserting this equality in Eq. (I8) we then obtain that the bound of Eq. (I7) is attained.

Summarizing, we have proven that the "best way to fail" is via a pure quantum operation. Iterating the same argument, we obtain that the optimal strategy at each step is described by a binary filter consisting of two pure quantum operations, with Kraus operators $B_{\mathrm{succ}}^{(k)}$ and $B_{\mathrm{fail}}^{(k)}$, respectively. Expanding the state at the kth step as

$$|\varphi^{(k)}\rangle = \sum_E \sqrt{p_E^{(k)}} |\varphi_E^{(k)}\rangle,$$

the successful Kraus operator is determined in an essentially unique way by Proposition 5, which yields the condition

$$B_{\mathrm{succ}}^{(k)} |\varphi_E^{(k)}\rangle = \left[\min_{E' \in \mathsf{Sp}(\varphi^{(k)}) \cap \mathsf{Sp}(\psi)} \sqrt{\frac{p_{E'}^{(k)}}{q_{E'}}} \right] \sqrt{\frac{q_E}{p_E^{(k)}}} |\psi_E\rangle \tag{I9}$$

for every energy E in $\mathsf{Sp}(\varphi^{(k)})$. The unsuccessful Kraus operator is then given by

$$B_{\mathrm{fail}}^{(k)} := \sqrt{I - B_{\mathrm{succ}}^{(k)\dagger} B_{\mathrm{succ}}^{(k)}}, \tag{I10}$$

and its definition is essentially unique, up to the application of an energy-preserving unitary on the output and to a possible redefinition of $B_{\mathrm{fail}}^{(k)}$ outside the relevant subspace.

Applying iteratively Eqs. (I9) and (I10) it is easy to obtain that the eigenstates $|\varphi_E^{(k)}\rangle$ are independent of k; i.e., one has

$$|\varphi_E^{(k)}\rangle \equiv |\varphi_E\rangle \quad \forall k = 1, \ldots, K, \forall E \in \mathsf{Sp}(\varphi^{(k)}).$$

This condition implies that Eq. (I9) can be rewritten as

$$B_{\mathrm{succ}}^{(k)}|\varphi_E\rangle = \left[\min_{E' \in \mathsf{Sp}(\varphi^{(k)}) \cap \mathsf{Sp}(\psi)} \sqrt{\frac{p_{E'}^{(k)}}{q_{E'}}} \right] \sqrt{\frac{q_E}{p_E^{(k)}}} |\psi_E\rangle$$

for every energy E in $\mathsf{Sp}(\varphi^{(k)})$.

APPENDIX J: KRAUS OPERATORS OF THE RECURSIVE PROTOCOL

Here we characterize the form of the successful Kraus operators M_k, with $k \leqslant K \leqslant L$. Physically, the operator M_k corresponds to the event that one succeeds at the kth round, after having failed in the first $k - 1$ rounds, namely

$$M_k = B_{\mathrm{succ}}^{(k)} B_{\mathrm{fail}}^{(k-1)} \cdots B_{\mathrm{fail}}^{(1)}. \tag{J1}$$

To characterize M_k we first analyze the operators $B_{\mathrm{fail}}^{(i)}$, with $i = 1, \ldots, k - 1$. Combining Eqs. (I9) and (I10), we obtain that $B_{\mathrm{fail}}^{(i)}$ satisfies the condition

$$B_{\mathrm{fail}}^{(i)}|\varphi_E\rangle = 0 \quad \forall E \in \mathsf{R}_1^{(i)}, \tag{J2}$$

where $\mathsf{R}_1^{(i)}$ is the set defined by

$$\mathsf{R}_1^{(i)} := \left\{ E \in \mathsf{Sp}(\varphi^{(i)}) \cap \mathsf{Sp}(\psi) \,\middle|\, \frac{p_E^{(i)}}{q_E} = r_1^{(i)} \right\}, \tag{J3}$$

$r_1^{(i)}$ being the minimum nonzero value of the ratio $r_E^{(i)} = p_E^{(i)}/q_E$. Now, the key observation is provided by the following:

Lemma 2. The set $\mathsf{R}_1^{(i)}$ coincides with the set R_i defined in Eq. (K1).

Proof. The proof is by recursion over i, based on the relation

$$\frac{p_E^{(i+1)}}{q_E} = \frac{\left\| P_E B_{\mathrm{fail}}^{(i)}|\varphi^{(i)}\rangle \right\|^2}{q_E \left\| B_{\mathrm{fail}}^{(i)}|\varphi^{(i)}\rangle \right\|^2}$$

$$= \frac{\langle \varphi^{(i)}| P_E (I - B_{\mathrm{succ}}^{(i)\dagger} B_{\mathrm{succ}}^{(i)}) P_E |\varphi^{(i)}\rangle}{q_E \left\| B_{\mathrm{fail}}^{(i)}|\varphi^{(i)}\rangle \right\|^2}$$

$$= \frac{p_E^{(i)} \left\{ 1 - \left[\min_{E' \in \mathsf{Sp}(\varphi^{(i)}) \cap \mathsf{Sp}(\psi)} \frac{p_{E'}^{(i)}}{q_{E'}} \right] \frac{q_E}{p_E^{(i)}} \right\}}{q_E \left\| B_{\mathrm{fail}}^{(i)}|\varphi^{(i)}\rangle \right\|^2}$$

$$= \left[\frac{p_E^{(i)}}{q_E} - \min_{E' \in \mathsf{Sp}(\varphi^{(i)}) \cap \mathsf{Sp}(\psi)} \frac{p_{E'}^{(i)}}{q_{E'}} \right] \left\| B_{\mathrm{fail}}^{(i)}|\varphi^{(i)}\rangle \right\|^{-2}, \tag{J4}$$

where the first equality follows from the definition $|\varphi^{(i+1)}\rangle := B_{\mathrm{fail}}^{(i)}|\varphi^{(i)}\rangle / \|B_{\mathrm{fail}}^{(i)}|\varphi^{(i)}\rangle\|$, the second from the relation $B_{\mathrm{fail}}^{(i)} := \sqrt{I - B_{\mathrm{succ}}^{(i)\dagger} B_{\mathrm{succ}}^{(i)}}$, and the third from Eq. (I9). Equation (J4) shows that the set of energies for which the ratio $p_E^{(i+1)}/q_E$ has the smallest value coincides with the set of energies for which the ratio $p_E^{(i)}/q_E$ has the second smallest value, which in turn coincides with the set of energies for which the ratio p_E/q_E

has the $(i + 1)$th smallest value. By definition of the sets $\mathsf{R}_1^{(i)}$ and R_i, this proves the thesis. ∎

Using the above lemma, Eq. (J2) becomes

$$B_{\mathrm{fail}}^{(i)}|\varphi_E\rangle = 0 \quad \forall E \in \mathsf{R}_i. \tag{J5}$$

We now use this relation to determine the form of the Kraus operator M_k. Setting $m_E^{(k)} := \|M_k|\varphi_E\rangle\|^2$, the definition of M_k [Eq. (J1)] gives the bound

$$m_E^{(k)} \leqslant \left\| B_{\mathrm{fail}}^{(i)}|\varphi_E\rangle \right\|^2 \quad \forall i = 1, \ldots, k - 1.$$

Combining this bound with Eq. (J5) we obtain the condition

$$m_E^{(k)} = 0 \quad \forall E \in \bigcup_{i=1}^{k-1} \mathsf{R}_i \equiv \mathsf{U}_{k-1}, \tag{J6}$$

which shows that the operator M_k annihilates the subspace spanned by the energy eigenstates with eigenvalues in U_{k-1}.

The action of M_k on the remaining eigenstates is determined by the fact that when M_k is decomposed as in Eq. (J1), the last operator acting in the sequence is $B_{\mathrm{succ}}^{(k)}$. Hence, we know that the initial amplitude $\sqrt{p_E}$ should be modulated to $\sqrt{q_E}$ for all the energy eigenvalues that survived the first $k - 1$ rounds; that is,

$$m_E^{(k)} = c_k \frac{q_E}{p_E} \quad \forall E \in \mathsf{Sp}(\varphi) \setminus \mathsf{U}_{k-1}, \tag{J7}$$

where $c_k > 0$ is a suitable proportionality constant. To determine c_k, note that the trace-preserving condition for the instrument $\{\mathcal{M}_k\}_{k=1}^{K+1}$ is equivalent to

$$\sum_{k=1}^{K+1} m_E^{(k)} = 1 \quad \forall E.$$

Combining this fact with Eqs. (J6) and (J7) we then obtain the recursion relations

$$c_1 \frac{q_E}{p_E} = 1 \quad \forall E \in \mathsf{R}_1,$$

$$(c_1 + c_2) \frac{q_E}{p_E} = 1 \quad \forall E \in \mathsf{R}_2,$$

$$\vdots$$

$$\left(\sum_{k=1}^{K} c_k \right) \frac{q_E}{p_E} = 1 \quad \forall E \in \mathsf{R}_K.$$

Finally, using the definition of the sets R_k [Eq. (K1)] we can solve the system of equations, obtaining $c_k = r_k - r_{k-1}$ for every $k \in \{1, \ldots, K\}$, having set $r_0 := 0$. In conclusion, the action of the successful Kraus operators is given by

$$M_k|\varphi_E\rangle = \begin{cases} 0 & E \in \mathsf{U}_{k-1}, \\ \sqrt{(r_k - r_{k-1}) \frac{q_E}{p_E}} |\psi_E\rangle & E \in \mathsf{Sp}(\varphi) \setminus \mathsf{U}_{k-1}. \end{cases} \tag{J8}$$

The expressions of the fidelity and of the probability of success, anticipated in Eqs. (55) and (56), can be easily derived from the above equation. Indeed, we have

$$F_{\max}^{(k)} = \frac{|\langle \psi | M_k | \varphi \rangle|^2}{\|M_k|\varphi\rangle\|^2} = \sum_{E \in \mathsf{Sp}(\varphi) \setminus \mathsf{U}_{k-1}} q_E$$

and

$$
p_{\mathrm{succ}}^{(k)} = \|M_k|\varphi\rangle\|^2
$$

$$
= (r_k - r_{k-1})\left(\sum_{E \in \mathsf{Sp}(\varphi)\setminus \mathsf{U}_{k-1}} q_E\right)
$$

$$
= (r_k - r_{k-1})F_{\max}^{(k)}.
$$

APPENDIX K: THE ULTIMATE LIMIT FOR PROBABILISTIC TRANSITIONS FROM MIXED TO PURE STATES

Proof of Theorem 4. Let $\mathcal{M}$ be an energy-preserving quantum channel that acts on the input state ρ. Its fidelity for the transition $\rho \to |\psi\rangle\langle\psi|$ is given by

$$
F = \langle\psi|\mathcal{M}(\rho)|\psi\rangle
$$

$$
= \sum_{E,E'} \sqrt{q_E q_{E'}}\langle\psi_E|\mathcal{M}(P_E \rho P_{E'})|\psi_{E'}\rangle
$$

$$
\leqslant \sum_{E,E'} \sqrt{q_E q_{E'}}|\langle\psi_E|\mathcal{M}(\rho_{E,E'})|\psi_{E'}\rangle|
$$

$$
\leqslant \sum_{E,E'} \sqrt{q_E q_{E'}}\|\mathcal{M}(\rho_{E,E'})\|_1,
$$

the last inequality following from the relation $|\operatorname{Tr}[AB]| \leqslant \|A\|_\infty \|B\|_1$ applied to the operators $A = |\psi_{E'}\rangle\langle\psi_E|$ and $B = \mathcal{M}(\rho_{E,E'})$. Since $\mathcal{M}$ is a trace-preserving quantum operation, we then obtain the bound

$$
F \leqslant \sum_{E,E'} \sqrt{q_E q_{E'}}\|\rho_{E,E'}\|_1.
$$

The bound can be achieved when the state ρ is block positive, namely when there exists an orthonormal basis $\{|\varphi_{E,k}\rangle | k = 1,\ldots,d_E\}$ for each eigenspace with energy E such that all matrices

$$
[\rho_{E,E'}] = [\langle\varphi_{E,k}|\rho|\varphi_{E',l}\rangle] \quad k,l \leqslant \min\{d_E, d_{E'}\}
$$

are (square and) positive semidefinite. Then, one can define the Kraus operators

$$
A_k = \sum_{E:k \leqslant d_E} |\psi_E\rangle\langle\varphi_{E,k}|
$$

and the quantum channel $\mathcal{A}(\cdot) = \sum_k A_k \cdot A_k^\dagger$. With this definition, one has

$$
F = \langle\psi|\mathcal{A}(\rho)|\psi\rangle
$$

$$
= \sum_k \sum_{E:k \leqslant d_E} \sum_{E':k \leqslant d_{E'}} \sqrt{q_E q_{E'}}\langle\varphi_{E,k}|\rho|\varphi_{E',k}\rangle
$$

$$
= \sum_{E,E'} \sum_{k \leqslant \min\{d_E,d_{E'}\}} \sqrt{q_E q_{E'}}\langle\varphi_{E,k}|\rho|\varphi_{E',k}\rangle
$$

$$
= \sum_{E,E'} \sqrt{q_E q_{E'}} \operatorname{Tr}[\rho_{E,E'}]
$$

$$
= \sum_{E,E'} \sqrt{q_E q_{E'}}\|\rho_{E,E'}\|_1,
$$

having used the fact that $\rho_{E,E'}$ is positive semidefinite, and therefore $\operatorname{Tr}[\rho_{E,E'}] = \|\rho_{E,E'}\|_1$. ∎

Proof of Theorem 5. Let $\mathcal{M}$ be an energy-preserving quantum channel and let M be its Choi operator, defined as

$$
M = (\mathcal{M} \otimes \mathcal{I})(|I\rangle\!\rangle\langle\!\langle I|)
$$

with $|I\rangle\!\rangle = \sum_n |n\rangle|n\rangle$ being the unnormalized maximally entangled state on $\mathcal{H}_1 \otimes \mathcal{H}_2$, $\mathcal{H}_1 \simeq \mathcal{H}_2 \simeq \mathcal{H}$. In the Choi representation, the energy-preserving condition is equivalent to the requirement

$$
\Pi_0 M \Pi_0 = M, \quad \Pi_0 = \bigoplus_E (P_E \otimes P_E).
$$

Now, the fidelity can be expressed as

$$
F = \frac{\langle\psi|\mathcal{M}(\rho)|\psi\rangle}{\operatorname{Tr}[\mathcal{M}(\rho)]}
$$

$$
= \frac{\operatorname{Tr}[M(|\psi\rangle\langle\psi| \otimes \rho^T)]}{\operatorname{Tr}[M(I \otimes \rho^T)]}
$$

$$
= \frac{\operatorname{Tr}[M(|\psi\rangle\langle\psi| \otimes \rho^T)]}{\operatorname{Tr}[\Pi_0 M \Pi_0 (I \otimes \rho^T)]}
$$

$$
= \operatorname{Tr}[\sigma R^{-1/2}(|\psi\rangle\langle\psi| \otimes \rho^T)R^{-1/2}],
$$

with

$$
R = \Pi_0 (I \otimes \rho^T)\Pi_0
$$

$$
= \bigoplus_E (P_E \otimes \rho_{E,E}^T) \tag{K1}
$$

and

$$
\sigma = \frac{\sqrt{R}M\sqrt{R}}{\operatorname{Tr}[MR]}.
$$

Since σ is a density matrix, we have the achievable upper bound

$$
F \leqslant \sup_{\sigma:\sigma \geqslant 0,\, \operatorname{Tr}[\sigma]=1} \operatorname{Tr}[\sigma R^{-1/2}(|\psi\rangle\langle\psi| \otimes \rho^T)R^{-1/2}]
$$

$$
= \|R^{-1/2}(|\psi\rangle\langle\psi| \otimes \rho^T)R^{-1/2}\|_\infty
$$

$$
= \left\|\sum_{E,E'} \sqrt{q_E q_{E'}}|\psi_E\rangle\langle\psi_{E'}| \otimes (\rho_{E,E}^{-\frac{1}{2}})^T \rho_{E,E'}^T (\rho_{E',E'}^{-\frac{1}{2}})^T\right\|_\infty
$$

$$
\equiv \|A\|_\infty,
$$

achieved if and only if the support of σ is contained in the eigenspace of A with maximum eigenvalue. Hence, we must have

$$
A\sigma = F_{\max}\sigma
$$

and

$$
M = \gamma R^{-1/2}\sigma R^{-1/2}, \tag{K2}
$$

for a suitable constant $\gamma > 0$. Note that γ is equal to the success probability; indeed, we have

$$
p_{\mathrm{succ}} = \operatorname{Tr}[M(I \otimes \rho^T)]
$$

$$
= \gamma \operatorname{Tr}[R^{-1/2}\sigma R^{-1/2}(I \otimes \rho^T)]
$$

$$
\begin{aligned}
&= \gamma \sum_{E,E'} \mathrm{Tr}\left[\left(P_E \otimes \rho_{E,E}^T\right)^{-1/2} \sigma \right.\\
&\quad\left. \times \left(P_{E'} \otimes \rho_{E',E'}^T\right)^{-1/2}(I \otimes \rho^T)\right]\\
&= \gamma \sum_E \mathrm{Tr}\left\{\sigma\left[P_E \otimes \left(\rho_{E,E}^T\right)^{-1/2}\rho^T\left(\rho_{E,E}^T\right)^{-1/2}\right]\right\}\\
&= \gamma \sum_E \mathrm{Tr}[\sigma(P_E \otimes Q_E)],
\end{aligned}
$$

where Q_E is the projector onto the support of $\rho_{E,E}$. Now, note that the support of A is contained in the support of the projector $P = \sum_E P_E \otimes Q_E$. Since the support of σ is contained in the support of A, we conclude

$$
p_{\text{succ}} = \gamma\, \mathrm{Tr}\left[\sigma\left(\sum_E P_E \otimes Q_E\right)\right] = \gamma.
$$

Finally, the maximum value of γ can be derived from the trace nonincreasing property of the quantum operation $\mathcal{M}$. In terms of the Choi operator, the nonincreasing property reads

$\mathrm{Tr}_1[M] \leqslant I$, where Tr_{out} denotes the trace on the output Hilbert space. Using Eqs. (K2) and (K1) we obtain the relation

$$
\begin{aligned}
\mathrm{Tr}_1[M] &= \gamma \sum_{E,E'} \mathrm{Tr}_1\left[\left(P_E \otimes \rho_{E,E}^T\right)^{-1/2}\sigma\left(P_{E'} \otimes \rho_{E',E'}^T\right)^{-1/2}\right]\\
&= \gamma \sum_E \left(\rho_{E,E}^T\right)^{-1/2}\sigma_E\left(\rho_{E,E}^T\right)^{-1/2},
\end{aligned}
$$

where $\sigma_E := \mathrm{Tr}_1[(P_E \otimes I)\sigma]$ are the unnormalized states on the input system, conditional on the outcomes of an energy measurement on the output system. Hence, the trace nonincreasing condition $\mathrm{Tr}_1[M] \leqslant I$ is guaranteed if and only if

$$
\gamma \leqslant \min_E \frac{1}{\left\|\left(\rho_{E,E}^T\right)^{-1/2}\sigma_E\left(\rho_{E,E}^T\right)^{-1/2}\right\|_\infty}.
$$

In conclusion, the maximum probability of success is given by

$$
p_{\text{succ}} = \max_\sigma \min_E \frac{1}{\left\|\left(\rho_{E,E}^T\right)^{-1/2}\sigma_E\left(\rho_{E,E}^T\right)^{-1/2}\right\|_\infty}. \qquad \blacksquare
$$

[1] S. Haroche, Rev. Mod. Phys. **85**, 1083 (2013).

[2] D. Wineland, Rev. Mod. Phys. **85**, 1103 (2013).

[3] I. Georgescu, S. Ashhab, and F. Nori, Rev. Mod. Phys. **86**, 153 (2014).

[4] M. Fuechsle, J. Miwa, S. Mahapatra, H. Ryu, S. Lee, O. Warschkow, L. Hollenberg, G. Klimeck, and M. Simmons, Nat. Nanotechnol. **7**, 242 (2012).

[5] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. White, J. Mutus, A. Fowler, B. Campbell *et al.*, Nature (London) **508**, 500 (2014).

[6] B. Hensen, H. Bernien, A. Dréau, A. Reiserer, N. Kalb, M. Blok, J. Ruitenberg, R. Vermeulen, R. Schouten, C. Abellán *et al.*, Nature (London) **526**, 682 (2015).

[7] G. Fettweis and E. Zimmermann, in *Proceedings of the 11th International Symposium on Wireless Personal Multimedia Communications* (Oulu Centre for Wireless Communications, 2008), Vol. 2, p. 6.

[8] A. M. Ionescu and H. Riel, Nature (London) **479**, 329 (2011).

[9] O. A. Mukhanov, IEEE Trans. Appl. Supercond. **21**, 760 (2011).

[10] A. Berl, E. Gelenbe, M. Di Girolamo, G. Giuliani, H. De Meer, M. Q. Dang, and K. Pentikousis, Comput. J. **53**, 1045 (2010).

[11] S. Mittal, Int. J. Comput. Aided Eng. Tech. **6**, 440 (2014).

[12] M. Aspelmeyer, T. Kippenberg, and F. Marquardt, Rev. Mod. Phys. **86**, 1391 (2014).

[13] P. Lodahl, S. Mahmoodian, and S. Stobbe, Rev. Mod. Phys. **87**, 347 (2015).

[14] Z. Xiang, S. Ashhab, J. You, and F. Nori, Rev. Mod. Phys. **85**, 623 (2013).

[15] P. Zhang, K. Aungskunsiri, E. Martín-López, J. Wabnig, M. Lobino, R. Nock, J. Munns, D. Bonneau, P. Jiang, H. Li *et al.*, Phys. Rev. Lett. **112**, 130501 (2014).

[16] J. Medford, J. Beil, J. Taylor, S. Bartlett, A. Doherty, E. Rashba, D. DiVincenzo, H. Lu, A. Gossard, and C. M. Marcus, Nat. Nanotechnol. **8**, 654 (2013).

[17] J. Fei, J.-T. Hung, T. S. Koh, Y.-P. Shim, S. N. Coppersmith, X. Hu, and M. Friesen, Phys. Rev. B **91**, 205434 (2015).

[18] C. Rigetti, A. Blais, and M. Devoret, Phys. Rev. Lett. **94**, 240502 (2005).

[19] Y. Aharonov and L. Vaidman, Phys. Rev. A **41**, 11 (1990).

[20] S. Aaronson, Proc. R. Soc. A **461**, 3473 (2005).

[21] Y. Sun, M. Hillery, and J. A. Bergou, Phys. Rev. A **64**, 022311 (2001).

[22] Y. Sun, J. A. Bergou, and M. Hillery, Phys. Rev. A **66**, 032315 (2002).

[23] G. Sentís, E. Bagan, J. Calsamiglia, and R. Muñoz-Tapia, Phys. Rev. A **88**, 052304 (2013).

[24] J. Fiurášek, New J. Phys. **8**, 192 (2006).

[25] B. Gendra, E. Ronco-Bonvehi, J. Calsamiglia, R. Muñoz-Tapia, and E. Bagan, Phys. Rev. A **88**, 012128 (2013).

[26] B. Gendra, E. Ronco-Bonvehi, J. Calsamiglia, R. Muñoz-Tapia, and E. Bagan, Phys. Rev. Lett. **110**, 100501 (2013).

[27] G. Chiribella, Y. Yang, and A. C.-C. Yao, Nat. Commun. **4**, 2915 (2013).

[28] G. Chiribella and Y. Yang, Front. Phys. **11**, 110304 (2016).

[29] J. Fiurášek, Phys. Rev. A **70**, 032308 (2004).

[30] T. Ralph and A. Lund, in *Quantum Communication Measurement and Computing: Proceedings of 9th International Conference*, edited by A. Lvovsky (AIP, Melville, 2009), pp. 155–160.

[31] G. Chiribella and J. Xie, Phys. Rev. Lett. **110**, 213602 (2013).

[32] X. Zhao and G. Chiribella, Phys. Rev. A **95**, 042303 (2017).

[33] F. Levi and F. Mintert, New J. Phys. **16**, 033007 (2014).

[34] T. Baumgratz, M. Cramer, and M. B. Plenio, Phys. Rev. Lett. **113**, 140401 (2014).

[35] J. A. Vaccaro, F. Anselmi, H. M. Wiseman, and K. Jacobs, Phys. Rev. A **77**, 032114 (2008).

[36] M. Lostaglio, K. Korzekwa, D. Jennings, and T. Rudolph, Phys. Rev. X **5**, 021001 (2015).

[37] J. Åberg, Phys. Rev. Lett. **113**, 150402 (2014).

[38] P. Ćwikliński, M. Studziński, M. Horodecki, and J. Oppenheim, Phys. Rev. Lett. **115**, 210403 (2015).

[39] I. Marvian and R. W. Spekkens, Phys. Rev. A **94**, 052324 (2016).

[40] J. I. Cirac, A. K. Ekert, and C. Macchiavello, Phys. Rev. Lett. **82**, 4344 (1999).

[41] U. L. Andersen, R. Filip, J. Fiurášek, V. Josse, and G. Leuchs, Phys. Rev. A **72**, 060301 (2005).

[42] G. M. D'Ariano, C. Macchiavello, and P. Perinotti, Phys. Rev. Lett. **95**, 060503 (2005).

[43] J. I. Cirac and P. Zoller, Phys. Rev. Lett. **74**, 4091 (1995).

[44] H. Häffner, C. F. Roos, and R. Blatt, Phys. Rep. **469**, 155 (2008).

[45] A. E. Siegman, IEEE J. Sel. Top. Quantum Electron. **6**, 1380 (2000).

[46] A. E. Siegman, IEEE J. Sel. Top. Quantum Electron. **6**, 1389 (2000).

[47] I. Bloch, Nat. Phys. **1**, 23 (2005).

[48] A. Derevianko and H. Katori, Rev. Mod. Phys. **83**, 331 (2011).

[49] T. J. Kippenberg and K. J. Vahala, Science **321**, 1172 (2008).

[50] F. Marquardt and S. M. Girvin, Physics **2**, 40 (2009).

[51] O. Bratteli and D. W. Robinson, *Operator Algebras and Quantum Statistical Mechanics 1: C∗- and W∗-Algebras. Symmetry Groups. Decomposition of States*, 2nd ed., Theoretical and Mathematical Physics (Springer, Berlin, 2003).

[52] G. Lindblad, Lett. Math. Phys. **47**, 189 (1999).

[53] M. Ozawa, J. Math. Phys. **25**, 79 (1984).

[54] E. P. Wigner, Z. Phys. **133**, 101 (1952).

[55] H. Araki and M. M. Yanase, Phys. Rev. **120**, 622 (1960).

[56] M. Ozawa, Phys. Rev. Lett. **88**, 050402 (2002).

[57] M. Ozawa, Phys. Rev. Lett. **89**, 057902 (2002).

[58] J. Gea-Banacloche and M. Ozawa, Phys. Rev. A **74**, 060301 (2006).

[59] I. Marvian and R. W. Spekkens, arXiv:1212.3378.

[60] M. Ahmadi, D. Jennings, and T. Rudolph, New J. Phys. **15**, 013057 (2013).

[61] M. Navascués and S. Popescu, Phys. Rev. Lett. **112**, 140502 (2014).

[62] P. Bush, M. Grabowski, and P. J. Lahti, *Operational Quantum Physics* (Springer, Berlin, 1997).

[63] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, Vol. 1 (Springer, Berlin, 2011).

[64] T. Heinosaari and M. Ziman, *The Mathematical Language of Quantum Theory: From Uncertainty to Entanglement* (Cambridge University Press, Cambridge, 2012).

[65] M. Partanen, T. Häyrynen, J. Oksanen, and J. Tulkki, Phys. Rev. A **86**, 063804 (2012).

[66] G. M. D'Ariano and C. Macchiavello, Phys. Rev. A **67**, 042306 (2003).

[67] G. Lüders, Ann. Phys. **443**, 322 (1950).

[68] P. Busch, P. J. Lahti, and P. Mittelstaedt, *The Quantum Theory of Measurement* (Springer, Berlin, 1991).

[69] J. Calsamiglia, Nat. Phys. **10**, 91 (2014).

[70] M. Ying and Y. Feng, Acta Informatica **47**, 221 (2010).

[71] D. P. DiVincenzo, Science **270**, 255 (1995).

[72] A. Paetznick and K. Svore, Quantum Inf. Comput. **14**, 1277 (2014).

[73] A. Bocharov, M. Roetteler, and K. Svore, Phys. Rev. Lett. **114**, 080502 (2015).

[74] N. Wiebe and M. Roetteler, Quantum Inf. Comput. **16**, 134 (2016).

[75] J. Anders, D. K. L. Oi, E. Kashefi, D. E. Browne, and E. Andersson, Phys. Rev. A **82**, 020301 (2010).

[76] M. Ozols, M. Roetteler, and J. Roland, ACM Transactions on Computation Theory (TOCT) **5**, 11 (2013).

[77] M. A. Nielsen and I. L. Chuang, Phys. Rev. Lett. **79**, 321 (1997).

[78] R. D. Sorkin, Mod. Phys. Lett. A **9**, 3119 (1994).

[79] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, London, 1976).

[80] V. Bužek, R. Derka, and S. Massar, Phys. Rev. Lett. **82**, 2207 (1999).

[81] D. W. Berry and H. M. Wiseman, Phys. Rev. Lett. **85**, 5098 (2000).

[82] V. Scarani, S. Iblisdir, N. Gisin, and A. Acín, Rev. Mod. Phys. **77**, 1225 (2005).

[83] N. J. Cerf and J. Fiurasek, Prog. Opt. **49**, 455 (2006).

[84] C. M. Caves, Phys. Rev. D **26**, 1817 (1982).

[85] S. Pandey, Z. Jiang, J. Combes, and C. M. Caves, Phys. Rev. A **88**, 033852 (2013).

[86] R. Namiki, Phys. Rev. A **83**, 042323 (2011).

[87] M. Gregoratti and R. F. Werner, J. Mod. Opt. **50**, 915 (2003).

[88] G. D'Ariano, S. Facchini, P. Perinotti, and M. Sacchi, Phys. Lett. A **373**, 3011 (2009).

[89] D. Oi, Eur. Phys. J. D **68**, 259 (2014).

[90] M. Owari, M. B. Plenio, E. S. Polzik, A. Serafini, and M. M. Wolf, New J. Phys. **10**, 113014 (2008).

[91] S. Bravy, D. P. DiVincenzo, R. Oliveira, and B. M. Terhal, Quantum Inf. Comput. **8**, 0361 (2008).

[92] S. Bravyi and B. Terhal, SIAM J. Comput. **39**, 1462 (2009).

[93] G. Burkard, R. H. Koch, and D. P. DiVincenzo, Phys. Rev. B **69**, 064503 (2004).

[94] D. M. Ceperley, Rev. Mod. Phys. **67**, 279 (1995).

[95] G. M. D'Ariano, C. Macchiavello, P. Perinotti, and M. F. Sacchi, Phys. Lett. A **268**, 241 (2000).

[96] G. Chiribella and Y. Yang, New J. Phys. **16**, 063005 (2014).

[97] G. Gour and R. W. Spekkens, New J. Phys. **10**, 033023 (2008).

[98] I. Marvian and R. W. Spekkens, New J. Phys. **15**, 033001 (2013).

[99] I. Marvian and R. W. Spekkens, Phys. Rev. A **90**, 062110 (2014).

[100] G. Chiribella, G. M. D'Ariano, and P. Perinotti, J. Math. Phys. **50**, 042101 (2009).

[101] V. Narasimhachar and G. Gour, Nat. Commun. **6**, 7689 (2015).

[102] C. King, K. Matsumoto, M. Nathanson, and M. B. Ruskai, Markov Proc. Related Fields **13**, 391 (2007).

[103] M. Ziman and V. Bužek, Phys. Rev. A **72**, 022110 (2005).

[104] F. Buscemi, G. Chiribella, and G. M. D'Ariano, Phys. Rev. Lett. **95**, 090501 (2005).

[105] F. Buscemi, G. Chiribella, and G. D'Ariano, Open Syst. Inf. Dynamics **14**, 53 (2007).

[106] K. Brádler, P. Hayden, D. Touchette, and M. M. Wilde, Phys. Rev. A **81**, 062312 (2010).

[107] M. M. Wilde, A. Winter, and D. Yang, Commun. Math. Phys. **331**, 593 (2014).

[108] S. D. Bartlett, T. Rudolph, R. W. Spekkens, and P. S. Turner, New J. Phys. **8**, 58 (2006).