Optimal energy-preserving conversions of quantum coherence

Yuxiang Yang and Giulio Chiribella
Center for Quantum Information, Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing 100084, China

Quantum mechanics owes its name to the fact that certain physical quantities, like the energy of a hydrogen atom or the spin of an electron, are discrete. But even more distinctive is the existence of coherent superpositions, which provide an invaluable resource for quantum information processing and quantum technologies. The characterization, quantification, and manipulation of this resource are currently the object of intense research and are expected to contribute to the design of new high-performance quantum devices. In this paper we address the search for the best evolution that converts a given quantum superposition of energy eigenstates into another without exchanging energy with the surrounding environment. We consider both deterministic and probabilistic evolutions, obtained by measuring the environment and postselecting a subset of the outcomes. In both cases, we characterize the process that maximizes the fidelity with the target superposition. This characterization is used to design a branching sequence of probabilistic filters that increase the probability of success while reaching maximum fidelity at each iteration. We then show that a coherent superposition of different histories generated by such branching allows one to construct efficient approximations of the optimal fidelity-probability tradeoff, via a technique dubbed coherent coarse-graining. The benefits of our construction are illustrated in a number of applications to phase estimation, quantum cloning, coherent state amplification, and ancilla-driven computation.

I. INTRODUCTION

Coherent superpositions of energy eigenstates are a precious resource in a vast number of arenas, ranging from quantum metrology [1–3] to computing [4–7] and even biological systems [8]. Quite naturally, great attention has been devoted to the characterization and quantification of this resource in a variety of contexts [9–14]. From the point of view of resources, it is natural to ask how well a given superposition can be transformed into another via a given set of free transformations, whose execution is regarded as inexpensive [15]. The basic requirement on free transformations is that they should not generate superpositions from incoherent mixtures of energy eigenstates. Once this requirement is satisfied, a variety of choices remain open: one can choose e.g. the full set of transformations that map mixtures of eigenstates into mixtures of eigenstates [13, 14] or, instead, the smaller class of transformations that are covariant under time evolution [9–12]. An even more selective class is the class of energy-preserving transformations, whose realization does not require exchange of energy between the system and the surrounding environment. In principle, all such transformations can be implemented without consuming external energy sources, letting the system interact with an ancilla that remains in the ground space all throughout the evolution.

In this paper we investigate how well a coherent superposition of energy eigenstates can be transformed into another via an energy-preserving evolution. We consider not only deterministic evolutions, but also probabilistic evolutions induced by measurements. In the probabilistic case we require that the energy be preserved on average, allowing individual transformations to modulate the relative amplitudes of different energy eigenstates. Leveraging on this fact, we achieve conversions that would not be possible deterministically, following a postselection approach that has found numerous applications in quantum information, including stochastic transformations of entangled states [16–18], unambiguous state discrimination [19–23] and estimation [24–27], noiseless amplification [28–31] and super-replication [27], postselected [32] and ancilla-driven computation [33]. For the conversion of coherent superpositions, we show that the consists of only two pure transformations, i.e., two transformations that cannot be further refined. All optimal transformations have a highly constrained form: for example, in the deterministic case they must be equivalent to unitary gates—a fact that is not necessarily true for the broader class of phase-covariant transformations [34].

After characterizing the structure of the optimal processes, we consider a different scenario, where the probability of success is not fixed a priori, but can be decided on the fly by the experimenter. In this scenario we develop a recursive protocol, consisting of a sequence of two-outcome measurements, which produces the best approximation of the target at each step. Subsequent iterations of the protocol lead to an increasing probability of success, in a way that is akin to repeat-until-success protocols [35–37]. We then introduce the operation of coherent coarse-graining, which allows one to join different outcomes of the recursive protocol into a single quantum operation, keeping the same probability of success but increasing the fidelity with the target. Coherent coarse-graining provides a canonical way to approximate the optimal trade-off curve between fidelity and success probability. As a general technique, it can be applied also to problems that are more general the energy-preserving conversions discussed in this paper—essentially including all conversions of pure states subject to a set of linear constraints. The recursive protocol and its coherent coarse-graining are illustrated in a number of concrete
examples, including phase estimation, quantum cloning, amplification of coherent states, and ancilla-driven computation.

The paper is organized as follows. Section II introduces the basic framework. In Section 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 Section IV and apply it to several tasks in quantum information processing (Section V), including parameter estimation (Subsection V A), state cloning (Subsection V B), coherent light amplification (Subsection V C) and ancilla-driven computation (Subsection V D). Finally, the conclusions are given in Section VI, which discusses related works and future applications.

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 deterministic energy-preserving evolutions, discussing its structure and its relation with the classes of incoherent and time-covariant channels, the two classes of free transformations considered in the resource theory of coherence. We then move to probabilistic operations, characterizing the stochastic evolutions induced by energy non-disturbing measurements on the environment.

A. Energy-preserving channels

Consider a quantum system interacting with the surrounding environment through the Hamiltonian $H_{\text{int}}$. As a result of the interaction, the system and the environment evolve jointly via the unitary operator $U = \exp{-i(H_{\text{sys}} + H_{\text{env}} + H_{\text{int}}/\hbar)}$, where $H_{\text{sys}}$ and $H_{\text{env}}$ are the Hamiltonians of the system and of the environment, respectively, and $\Delta t$ is the interaction time. Suppose that the joint evolution preserves the energy of the system, namely

$$U^\dagger H_{\text{sys}} U = H_{\text{sys}}. \tag{1}$$

We refer to unitaries with this property as energy-preserving.

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

$$\mathcal{M}(\rho) = \text{Tr}_{\text{env}} \left[ U (\rho \otimes \rho_{\text{env}}) U^\dagger \right] \tag{2}$$

where $\text{Tr}_{\text{env}}$ denotes the partial trace over the Hilbert space of the environment. Clearly, the channel $\mathcal{M}$ preserves the expectation value of the energy. Even more strongly, the condition in Eq. (1) implies the relation

$$\mathcal{M}^\dagger (H_{\text{sys}}) = H_{\text{sys}}, \tag{3}$$

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

$$\mathcal{M}^\dagger (A) = \text{Tr}_{\text{env}} \left[ (I_{\text{sys}} \otimes \rho_{\text{env}}) U^\dagger (A \otimes I_{\text{env}}) U \right], \tag{4}$$

for arbitrary operators $A$. When Eq. (3) is satisfied, we say that $\mathcal{M}$ is an energy-preserving channel.

Eq. (3) implies that $\mathcal{M}$ does not preserve only the expectation value of the energy, but also its variance and all the higher order momenta, namely $\mathcal{M}^\dagger (H^n_{\text{sys}}) = H^n_{\text{sys}}$ for every integer $n$. This condition is equivalent to the requirement that the effective evolution $\mathcal{M}$ does not affect the probability distribution of the energy: for every state $\rho$ and for every eigenvalue $E$, one has the equality of probabilities

$$\text{Tr}[P_E \mathcal{M}(\rho)] = \text{Tr}[P_E \rho], \tag{5}$$

where $P_E$ is the projector on the eigenspace corresponding to $E$.

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. (3). The converse is also true: given an energy-preserving channel $\mathcal{M}$, one can always construct an interaction with an ancilla that does not change the energy of the system. To establish this fact, note that the map $\mathcal{M}^\dagger$ satisfies the condition

$$\mathcal{M}^\dagger (A) = A \quad \forall A \in \mathcal{A}, \tag{6}$$

where $\mathcal{A}$ is the Abelian algebra generated by the powers of the Hamiltonian. Note that the algebra $\mathcal{A}$ contains the identity and is closed under adjoint. Technically, algebras of this kind are known as unital $*$-algebras [38]. For a given algebra $\mathcal{A}$, the maps that satisfy Eq. (6) are characterized by a simple lemma due to Lindblad:

Lemma 1 (Lindblad [39]). 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

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

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

In the case of the Abelian algebra generated by the powers of $H_{\text{sys}}$, the commutation condition reduces to $[M_k, H_{\text{sys}}] = 0$, meaning that each Kraus operator $M_k$ must be of the block diagonal form

$$M_k = \bigoplus_E P_E M_k P_E, \tag{7}$$
with the sum running over spectrum of $H_{\text{sys}}$, which we will assume to be discrete all throughout the paper.

As a consequence of the block diagonal form (7), the isometry $V : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}_{\text{env}}$ defined by

$$V = \sum_{k=1}^{K} M_k \otimes |\phi_k\rangle,$$

(8)

where \{ |\phi_k\rangle \}_{k=1}^{K} are orthonormal states, satisfies the relation

$$V H_{\text{sys}} = (H_{\text{sys}} \otimes I_{\text{env}}) V,$$

(9)

which follows from the definition of $V$ in Eq. (8) and from the block diagonal form of Eq. (7). In turn, Eq. (9) implies that the isometry $V$ can be expressed as

$$V = U (I_{\text{sys}} \otimes |\phi_0\rangle)$$

(10)

where $|\phi_0\rangle \in \mathcal{H}_{\text{env}}$ is a pure state of the environment and $U$ is a unitary gate satisfying Eq. (1).

Summarizing, we have established the following:

**Proposition 1.** A quantum channel $\mathcal{M}$ is energy-preserving if and only if it can be realized via an interaction with the environment that preserves the energy of the system, namely $\mathcal{M}(\rho) = \text{Tr}_{\text{env}}[U (\rho \otimes |\phi_0\rangle \langle \phi_0|) U^\dagger]$, with $U$ satisfying the condition $U^\dagger H_{\text{sys}} U = H_{\text{sys}}$.

In fact, we will soon see that the unitary realization of energy-preserving channel satisfies an even stronger condition, namely that the total energy of the system and the environment is preserved and the environment remains in an eigenstate of the energy from the beginning to the end of the evolution.

**C. Relation to covariant channels**

We now highlight the relation between energy-preserving channels and channels that are covariant under time evolution [40], i.e. channels that satisfy the relation

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

(11)

Covariant channels are at the basis of the resource theory of asymmetry [9-11] and provide a possible candidate for the class of free operations in the resource theory of coherence. Every covariant channel can be realized via a unitary evolution—as in Eq. (2)—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:

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

(12)

(see e.g. [41]). 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.

It is easy to see that every energy-preserving channel is covariant, while the converse is not true in general. Physically, the difference can be understood in terms of the unitary realization: for a covariant channel the unitary $U$ preserves the sum of $H_{\text{sys}} + H_{\text{env}}$, while for an energy-preserving channel it preserves $H_{\text{sys}}$ and $H_{\text{env}}$ individually, as it follows from Eq. (12) and from proposition 1. This observation provides a strengthening of proposition 1, stating that every energy-preserving channel can be implemented via an interaction that causes no exchange of energy between the system and 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 via an interaction with the environment $\mathcal{M}(\rho) = \text{Tr}_{\text{env}}[U (\rho \otimes |\phi_0\rangle \langle \phi_0|) U^\dagger]$ such that $|\phi_0\rangle$ is an eigenstate of the energy and $U$ satisfies the conditions $U^\dagger H_{\text{sys}} U = H_{\text{sys}}$ and $U^\dagger H_{\text{env}} U = H_{\text{env}}$.

In particular, since the environment starts off in an energy eigenstate and the evolution preserves the energy, the state of the environment will remain in the same eigenspace at all times. This is the case, e.g. when the environment is a system of ultra-cold atoms, with a large energy gap between the ground space and the first excited states that effectively confines the evolution to the ground space. Generally, energy-preserving evolutions can be implemented by initializing the environment in a pure energy eigenstate and eventually returning it in a mixture of eigenstates with the same energy. In this sense, the evolution can be implemented at zero energy cost, at the price of an entropy increase in the environment.

**D. Relation to Hadamard channels, incoherent channels, and decoherence maps**

In the case of a Hamiltonian $H_{\text{sys}}$ with non-degenerate spectrum, the energy-preserving channels have been characterized in Ref. [42]: 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

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

(13)

where $C = [C_{mn}]$ is a positive matrix with diagonal elements equal to one. Channels of this are also known as Hadamard channels [43] in the literature on quantum Shannon theory, where they represent one of the important classes of channels with tractable capacity regions [44, 45]. Energy-preserving channels with non-degenerate 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 non-degenerate Hamiltonian, the energy-preserving channels are a special subset of incoherent channels [13, 14], 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 non-resourceful ones. On the contrary, typically they reduce quantum coherence, by damping down the off-diagonal elements of the density matrix [42]. For this reason, they have also been called decoherence maps in Refs. [42, 46, 47]. The inclusion relations among energy-preserving, covariant and incoherent operations are illustrated in Fig. 1.

E. Energy-preserving measurements

In addition to deterministic evolutions, we consider also the probabilistic evolutions induced by quantum measurements on the environment. Following Ozawa’s model of the measurement process [48], we suppose that, immediately after the interaction is switched off, the experimenter measures a “meter observable” \( M \) on the environment. In tune with the energy-preserving paradigm, we require that

1. the meter observable and the energy of the environment are compatible observables (\( [M, H_{\text{env}}] = 0 \))

2. the interaction preserves the energy of the system (\( U^\dagger H_{\text{sys}} U = H_{\text{sys}} \)).

In particular, these conditions are satisfied in the scenario of the Wigner-Araki-Yanase theorem [49–51], where condition 1 is known as Yanase’s condition [51].

According to quantum measurement theory [40, 48, 52, 53], the measurement of \( M \) induces a stochastic evolution of the state of the system, described by a quantum instrument, namely a collection of quantum operations (completely positive trace non-increasing maps) \( \{ \mathcal{M}_x \}_{x \in \mathcal{X}} \) subject to the normalization condition

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

for every quantum state \( \rho \). For a system prepared in the state \( \rho \), the measurement generates the outcome \( x \in \mathcal{X} \) with probability

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

and, conditionally on outcome \( x \), returns the system in the state

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

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

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

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

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

where \( \mathcal{M} \) is an energy-preserving channel. This observation motivates the following

Definition 1. We say that the instrument \( \{ \mathcal{M}_x \} \) is energy-preserving if the corresponding channel \( \mathcal{M} \) is energy-preserving.

It is not hard to see that every energy-preserving instrument arises from a measurement scheme as in Eq. (17), where the joint evolution \( U \) preserves \( H_{\text{sys}} \) and the observable \( M \) commutes with \( H_{\text{env}} \). This fact is illustrated 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
\[
U|0\rangle|0\rangle = |0\rangle|0\rangle,
\]
\[
U|1\rangle|0\rangle = |1\rangle|0\rangle,
\]
\[
U|0\rangle|1\rangle = 0|0\rangle|1\rangle,
\]
\[
U|1\rangle|1\rangle = |1\rangle|0\rangle|1\rangle.
\]

By measuring the meter observable $M = |1\rangle\langle 1| \otimes |0\rangle\langle 0|\}$ on the environment, one then obtain the instrument $\{M_0, M_1\}$ as effective evolution of the system.

It is worth stressing that, despite 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 $|1\rangle$, by discarding the cases where the measurement gives outcome $x = 0$. 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 $|+\rangle$, and the transformation will take place at no energy cost, as discussed in Example 1. However, this fact is not in contradiction with the conservation of the energy, which still holds on average over the two possible outcomes of the measurement.

Quantum measurements can be used to postselect a desired subset of evolutions. In this case, we refer to them as quantum filters. Mathematically, a filter is described by an instrument $\{M_x\}_{x \in X}$ along with a partition of outcome set $X$ into two disjoint subsets, $X_{\text{succ}}$ and $X_{\text{fail}}$, which correspond to successful and unsuccessful instances, respectively.

F. Relation to covariant instruments

Energy-preserving instruments are a special case of instruments that are covariant under time evolution [40], i.e. instruments $\{M_x\}_{x \in X}$ such that every quantum operation $M_x$ satisfies the condition (11). More specifically, the relation between covariant and energy-preserving instruments is expressed by the following

**Proposition 2** ([27]). Every covariant instrument $\{N_x\}_{x \in X}$ transforming states on $\mathcal{H}_{\text{in}}$ into states on $\mathcal{H}_{\text{out}}$ can be decomposed as
\[
N_x = C_x M_x \quad \forall x \in X
\]
where $\{M_x\}_{x \in X}$ is an energy-preserving instrument acting on the input and $C_x$ is a covariant channel transforming the input into the output system.

This result provides an additional motivation to the study of energy-preserving filters. Indeed, there is a large class of tasks where the optimal quantum strategy is covariant—this is the case, e.g. of phase-covariant cloning [27, 28, 34], phase-insensitive amplification [54], and phase-estimation [40, 55]. Proposition 2 establishes that energy-preserving filters are the canonical probabilistic element in all these tasks. The search for the optimal quantum operation is then split into two subproblems: i) the search for the optimal energy-preserving filter and ii) the search for the optimal deterministic operation. Problem i) coincides with the key question of this paper, about the optimal energy-preserving transformations of quantum states. This question will be addressed in the next Section.

III. OPTIMAL ENERGY-PRESERVING MANIPULATIONS OF QUANTUM COHERENCE

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 first formalize the problem and then address the optimality question, providing the general form of the optimal energy-preserving operation.

A. A state transformation game

Consider two coherent superpositions of energy eigenstates, denoted by $|\varphi\rangle$ and $|\psi\rangle$ and imagine that Alice and Bob play the following state transformation game: Alice prepares the system in the state $|\varphi\rangle$ and handles it to Bob, asking him to produce the state $|\psi\rangle$. Bob knows the states $|\varphi\rangle$ and $|\psi\rangle$, but in the implementation of his task is constrained by the fact that his operations should preserve the energy. After Bob has returned the system to Alice, she will check if the state is $|\psi\rangle$ or not using the binary measurement $\{P_0, P_1\}$, with $P_1 = |\psi\rangle\langle\psi|$ and $P_0 = I - |\psi\rangle\langle\psi|$. If the outcome is 1, she will assign score +1, otherwise she will assign score 0. This state transformation game is similar in spirit the task of quantum rejection sampling [56], wherein the goal is to generate a target superposition $|\psi\rangle$ using a black box $U_\varphi$ that prepares another superposition $|\varphi\rangle$ from a fixed state $|0\rangle$. Note however that having at disposal a coherent superposition $|\varphi\rangle$ is a weaker resource than being able to use the corresponding gate $U_\psi$, due to Nielsen and Chuang’s no-programming theorem [57].

Now, suppose that Bob is allowed to use a filter and to pass whenever the outcome is unsuccessful. Averaging over the successful instances, Bob’s expected payoff will be
\[
F = \sum_{x \in X_{\text{succ}}} p(x|\text{succ}) \langle\psi|\rho'_x|\psi\rangle,
\]

where $p(x|\text{succ})$ is the conditional probability of obtaining $x$ given that one successful outcome occurred and $\rho'_x$ is the output state conditional to the outcome $x$, given by Eq. (16) with $p = |\varphi\rangle\langle\varphi|$. Note that the payoff is nothing but the fidelity between the target state $|\psi\rangle$ and Bob’s
output conditional to the success of the filter. Making the filter explicit, the average payoff can be rewritten as

\[ F = \frac{\langle \psi | M_{\text{succ}} | \varphi \rangle \langle \varphi | \psi \rangle}{p_{\text{succ}}} \]

where \( M_{\text{succ}} \) is the CP map defined by

\[ M_{\text{succ}} := \sum_{x \in X_{\text{succ}}} M_x \]

and \( p_{\text{succ}} \) is the probability of success, given by

\[ p_{\text{succ}} = \text{Tr}[M_{\text{succ}} (|\varphi\rangle \langle \varphi|)] \]  

(21)

In a realistic situation, Bob will be interested not only in maximizing his payoff, but also in succeeding with sufficiently high probability. Requiring the success probability to be larger than a given threshold, the problem then becomes to find the energy-preserving quantum operation \( M_{\text{succ}} \) that maximizes the fidelity.

B. Optimality of pure binary filters

In the search of the optimal filter, it is immediately clear that one can restrict the attention to filters with two outcomes. This is because i) both the fidelity and the success probability depend only of the map \( M_{\text{succ}} \) defined in Eq. (20) and ii) if the original filter \( \{ M_x \} \) preserves the energy, so does the binary filter \( \{ M_{\text{succ}}, M_{\text{fail}} \} \) with \( M_{\text{fail}} := \sum_{x \in X_{\text{fail}}} M_x \).

We now show that the optimal filter can be chosen to be pure without loss of generality, meaning that the quantum operations \( M_x \) can be chosen of the form \( M_x(\cdot) = M_x \cdot M_x^\dagger \) for suitable operators \( M_x, x \in \{ \text{succ, fail} \} \).

**Theorem 2** (Optimality of pure filters). For every fixed value of the success probability, the energy-preserving filter that maximizes the fidelity is pure.

**Proof.** Let us expand the state \( |\varphi\rangle \) as

\[ |\varphi\rangle = \sum_E \sqrt{p_E} |\varphi_E\rangle \quad |\varphi_E\rangle := \frac{P_E |\varphi\rangle}{\|P_E |\varphi\rangle\|} \]

where \( P_E \) the projector on the energy eigenspace with eigenvalue \( E \) and \( p_E = \|P_E |\varphi\rangle\|^2 \). Similarly, we expand \( |\psi\rangle \) as \( |\psi\rangle = \sum_E \sqrt{q_E} |\psi_E\rangle \).

Consider a generic energy-preserving binary filter \( \{ M_{\text{succ}}, M_{\text{fail}} \} \), and expand the transformation \( M_{\text{succ}} \) in the Kraus form \( M_{\text{succ}}(\cdot) = \sum_k M_k \cdot M_k^\dagger \). By the energy-preserving condition [Eq. (7)] we have

\[ M_k = \bigoplus_E M_E^{(k)} \]

with \( M_E^{(k)} := P_E M_k P_E \). Fixing the probability of success to be \( p_{\text{succ}} \), we have the upper bound

\[ F \cdot p_{\text{succ}} = \sum_k \left( \langle \psi | M^{(k)} | \varphi \rangle \right)^2 \]

\[ = \sum_k \left( \sum_E \sqrt{p_E q_E} \langle \psi_E | M_E^{(k)} | \varphi_E \rangle \right)^2 \]

\[ \leq \sum_k \left( \sum_E \sqrt{p_E q_E} \| \langle \psi_E | M_E^{(k)} | \varphi_E \rangle \| \right)^2 \]

\[ \leq \sum_k \left( \sum_E \sqrt{p_E q_E} \| M_E^{(k)} | \varphi_E \rangle \| \right)^2 \]

Now, let us define the operator

\[ M_E := \sqrt{\sum_k M_E^{(k)} M_E^{(k)\dagger}} \]

(22)

and let us set \( a_E := \sqrt{\| M_E | \varphi_E \rangle \|} \) and \( b_E,k := \| M_E^{(k)} | \varphi_E \rangle \| / a_E \). With this notation, the upper bound becomes

\[ F \cdot p_{\text{succ}} \leq \sum_k \left( \sum_E \sqrt{p_E q_E} \frac{a_E}{b_E,k} b_E,k \right)^2 \]

\[ \leq \sum_k \left( \sum_E \sqrt{p_E q_E} \frac{a_E^2}{b_E,k} \right) \left( \sum_{E'} \sqrt{p_{E'} q_{E'}} b_{E',k}^2 \right) \]

\[ = \left( \sum_E \sqrt{p_E q_E} a_E^2 \right)^2 \]

\[ = \left( \sum_E \sqrt{p_E q_E} \| M_E | \varphi_E \rangle \| \right)^2 \]

(23)

The upper bound can be achieved by choosing a filter with successful Kraus operation \( M_{\text{succ}}(\cdot) = M' \cdot M'^\dagger \) defined by

\[ M' = \bigoplus_E M_E' \quad M_E' := U_E M_E \]

(24)

where \( U_E \) is any unitary operator on \( \mathcal{H}_E = P_E \mathcal{H} \) that transforms \( M_E | \varphi_E \rangle / \| M_E | \varphi_E \rangle \| \) into \( |\psi_E\rangle \). It is straightforward to prove that \( \| M_E | \varphi_E \rangle \|^2 = \sum_k \| M_E^{(k)} | \varphi_E \rangle \|^2 \), which means that the probability of success for the filter \( M_{\text{succ}}' \) is equal to the probability of success of \( M_{\text{succ}} \).}

**Theorem 2** is provides a simple characterization of the optimal transformation in the deterministic case, provided by the following
Corollary 1. For $p_{\text{succ}} = 1$, the optimal energy-preserving transformation is unitary and the optimal fidelity is given by

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

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 suboptimal for cloning qubit states on the equator of the Bloch sphere whenever the difference between the number of input copies and the number of clones is odd [34].

Theorem 2 provides an explicit expression also for the probabilistic transformation that attains the absolute maximum of the fidelity:

Corollary 2. For finite dimension Hilbert space $\mathcal{H} \simeq \mathbb{C}^d$, the maximum of the fidelity over all quantum operations is

$$F_{\text{max}} = \sum_{E \in \text{Sp}(\varphi)} q_E ,$$

where $\text{Sp}(\varphi)$ denotes the energy spectrum of the state $|\varphi\rangle$, defined as

$$\text{Sp}(\varphi) := \{ E \mid \langle \varphi | P_E | \varphi \rangle \neq 0 \} .$$

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

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

The quantum operation achieving maximum fidelity with maximum probability is pure and its Kraus operator satisfies the condition

$$M|\varphi_E\rangle = \left[ \min_{E \in \text{Sp}(\varphi) \cap \text{Sp}(\psi)} \sqrt{\frac{p_E}{q_E}} \right] \sqrt{\frac{q_E}{p_E}} |\psi_E\rangle$$

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

Proof. The bound $F \leq F_{\text{max}}$ follows immediately from Eq. (23) using Schwarz inequality. In order to achieve the bound, the Kraus operator $M$ must satisfy the condition

$$M|\varphi_E\rangle = c \sqrt{\frac{q_E}{p_E}} |\psi_E\rangle \quad \forall E \in \text{Sp}(\varphi) ,$$

(25)

where $c$ is a suitable constant. By the trace non-increasing condition, the constant $c$ must satisfy the relation $|c| \leq \sqrt{\frac{p_E}{q_E}}$ for every $E \in \text{Sp}(\varphi) \cap \text{Sp}(\psi)$. Now, the probability of success is given by

$$p_{\text{succ}} = |c|^2 \sum_{E \in \text{Sp}(\varphi)} q_E .$$

Clearly, the success probability is maximized by choosing $|c| = \min_{E \in \text{Sp}(\varphi) \cap \text{Sp}(\psi)} \sqrt{\frac{p_E}{q_E}}$, which also determines the form of the optimal Kraus operator $M$. \hfill \blacksquare

C. Coherent coarse-graining of a filter

The construction used in the proof of Theorem 2 is important in its own right: essentially, given an arbitrary instrument $\{M_x\}_{x \in X}$ and a partition of the outcomes into successful and unsuccessful ones, the construction of Eq. (24) provides a way to increase the fidelity by turning the set of transformations $\{M_x\}_{x \in \text{succ}}$ into a single, pure transformation $M'_{\text{succ}}$. We call $M'_{\text{succ}}$ the coherent coarse-graining of $M_{\text{succ}}$. The coherent coarse-graining can be visualized as a curve that splits the probability distribution $\{p_E\}$ into two subnormalized distributions, namely $\{p_E||M_E|\psi_E\rangle|^2\}$ and its complement.

A physical interpretation of coherent coarse-graining is as follows: for a generic instrument $\{M_x\}_{x \in X}$ one can define a POVM $\{P_x\}$ via the relation

$$P_x := M_x^\dagger (I) \quad \forall x \in X .$$

The coherent coarse-graining of $\{M_x\}_{x \in X}$ is then the pure instrument $\{M'_x\}_{x \in X}$ defined by

$$M'_x(\rho) := \sqrt{P_x} \rho \sqrt{P_x} \quad \forall x \in X .$$

When the POVM $\{P_x\}$ consists of projectors, this is exactly the state-reduction rule postulated by Lüders [58] based on ideas of minimal disturbance. In general, the coherently coarse-grained instrument represents the least disturbing state reduction associated to the POVM $\{P_x\}$ [59].

D. General form of the optimal filter

Theorem 2 guarantees that the optimal filter consists of pure transformations. In fact, it says even more, providing a stringent condition on the form of the optimal transformation. Let us consider the problem of maximizing the fidelity for a given value of the success probability. By Theorem 2, it is sufficient to consider pure transformations of the form $M_{\text{succ}}(\cdot) := M \cdot M^\dagger$, with $M = \bigoplus E M_E$ and

$$M_E|\varphi_E\rangle \propto |\psi_E\rangle \quad \forall E ,$$

Now, define the coefficients

$$m_E := ||M_E|\varphi_E\rangle ||^2$$

which have values in the interval $[0,1]$ for every energy $E$. With this definition, the probability of success (21) and the fidelity (19) can be expressed as

$$p_{\text{succ}} = \sum_E p_E m_E$$

and

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

(26)
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

\[
m_E = \begin{cases} 1 & P_{\text{succ}} - P_0 - q_E \frac{E}{p_E} \in S_0 \\ 1 - Q_0 & P_{\text{succ}} - P_0 - q_E \frac{E}{p_E} \in S_1 \end{cases}
\] (27)

where \( P_0 := \sum_{E \in S_0} P_E \) and \( Q_0 := \sum_{E \in S_0} q_E \). In other words, the optimal filter is completely determined by the choice of the set \( S_0 \). Inserting Eq. (27) into Eq. (26), the maximization of the fidelity is reduced to the maximization of the quantity

\[
\Omega[S_0] = \sum_{E \in S_0} \sqrt{P_E q_E} + \sqrt{(P_{\text{succ}} - P_0)(1 - Q_0)}.
\]

Examples of optimal filters of the form of Eq. (27) can be found in Ref. [26], which focused on the specific problem of phase-estimation. The application of energy-preserving filters to phase estimation will be discussed in more detail in subsection V A.

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 taylor-made for 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.

A. The optimal recursive protocol

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

1. For \( k = 1, \ldots, K - 1 \) apply the \( k \)-th filter
2. If the outcome is \( x = \text{succ} \), then terminate
3. If the outcome is \( x = \text{fail} \), then apply the \( (k + 1) \)-th filter.

Recursive protocols of this form are an example of the “quantum loop programs” studied in Ref. [60]. All these protocols can be can be visualized as decision trees of the following form

```
Filter 1
  \[\text{succ}\]
  \[\text{fail}\]
Filter 2
  \[\text{succ}\]
  \[\text{fail}\]
  Filter \( K \)
  \[\text{succ}\]
  \[\text{fail}\]
```

Protocols of the above form have been employed for different purposes, including entanglement concentration [61], implementation of quantum gates [35–37] and ancilla-driven computation [33].

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 at each individual round.

We consider the conversion problem \( |\varphi\rangle \rightarrow |\psi\rangle \) in the case of states \( |\varphi\rangle \) and \( |\psi\rangle \) in a finite dimensional Hilbert space \( \mathcal{H} \cong \mathbb{C}^d \), \( d < \infty \), or, more generally, states whose energy spectra intersect on a finite set of points, with \( |\text{Sp}(\varphi) \cap \text{Sp}(\psi)| \leq d \). For the optimal protocol we make a list of desiderata: for every \( k \in \{1, \ldots, K - 1\} \)

1. at the \( k \)-th round, the successful quantum operation should transform the input state \( \rho^{(k)} \) into the target \(|\psi\rangle\) with maximum fidelity
2. the optimal conversion \( \rho^{(k)} \rightarrow |\psi\rangle\langle\psi| \) must be achieved with maximum probability of success
3. conditional on the fulfillment of the first two requirements, the unsuccessful quantum operation at the \( k \)-th round should produce the state \( \rho^{(k+1)} \) that leads to maximum fidelity for the transformation \( \rho^{(k+1)} \rightarrow |\psi\rangle\langle\psi| \) 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 A. In the following we present the final result of the optimization and discuss the implications of this result.

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

\[
|\varphi^{(k)}\rangle = \bigoplus_{E} \sqrt{P_E^{(k)}} |\varphi_E\rangle,
\]
where the energy eigenstates are the same as in the decomposition of \( |\varphi\rangle \). With this notation, the successful quantum operation determined in an essentially unique way by the condition

\[
B^{(k)}_{\text{suc}} |\varphi_E\rangle = \min_{E' \in \text{Sp}(\varphi^{(k)}) \cap \text{Sp}(\psi)} \sqrt{\frac{p^{(k)}_E}{q^{(k)}_E}} \sqrt{q^{(k)}_E} |\psi_E\rangle \quad \forall E \in \text{Sp}(\varphi^{(k)}). \quad (28)
\]

Here the only freedom is in the way \( B^{(k)}_{\text{suc}} \) is defined in the subspace spanned by energy eigenstates outside the spectrum of \( |\varphi^{(k)}\rangle \). The form of Eq. (28) follows directly from the requirements 1 and 2 in our list of desiderata (cf. corollary 2). Similarly, the unsuccessful quantum operation is determined in an essentially unique way by the condition

\[
B^{(k)}_{\text{fail}} = \sqrt{I - B^{(k)}_{\text{suc}} B^{(k)}_{\text{suc}}^\dagger}, \quad (29)
\]

the only freedom being in the application of energy-preserving unitaries on the output and in the definition of \( B^{(k)}_{\text{fail}} \) in the subspace spanned by energy eigenstates outside the spectrum of \( |\varphi^{(k)}\rangle \). The form of Eq. (29) follows from the requirement 3 in our list. Remarkably, the quantum operation \( B^{(k)}_{\text{fail}} \) 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 equations (28) and (29). 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 introduce some notation. For every energy \( E \in \text{Sp}(\varphi) \cap \text{Sp}(\psi) \), we define the ratio \( r_E := q_E / p_E \) and arrange its possible values 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

\[
L \leq |\text{Sp}(\varphi) \cap \text{Sp}(\psi)| \leq d.
\]

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

\[
R_i := \left\{ E \in \text{Sp}(\varphi) \cap \text{Sp}(\psi) \mid \frac{p_E}{q_E} = r_i \right\} \quad (30)
\]

and we denote the union of the first \( k \) sets by

\[
U_k := \bigcup_{i=1}^k R_i. \quad (31)
\]

With this definition, the fidelity and the success probability at the \( k \)-th step can be expressed as

\[
F^{(k)}_{\text{max}} = \sum_{E \in \text{Sp}(\varphi) \setminus U_{k-1}} q_E. \quad (32)
\]

and

\[
p^{(k)}_{\text{suc}} = (r_k - r_{k-1}) \cdot F^{(k)}_{\text{max}}, \quad (33)
\]

respectively. Both equations can be proven easily using the characterization of the recursive protocol given in the next subsection. For the moment, let us examine their consequences. First, we 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. (32) and (33) turn out to be very useful for studying the tradeoff between fidelity and success probability. Indeed, they allow one to evaluate the probability that the protocol succeeds in one of the first \( T \) rounds, given by

\[
p_{\text{suc}}(T) := \sum_{k=1}^T p^{(k)}_{\text{suc}} = \sum_{E \in U_{T-1}} p_E + r_T F^{(T)}_{\text{max}}, \quad (34)
\]

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

\[
F(T) := \frac{\sum_{k=1}^T p^{(k)}_{\text{suc}}}{p_{\text{suc}}(T)} F^{(k)}_{\text{max}}. \quad (35)
\]

The tradeoff curve between \( F(T) \) and \( p_{\text{suc}}(T) \) will be illustrated in section 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 important to know what states are produced at every step of the protocol. Assuming that the total number of rounds is upper bounded as \( L \leq L \), the explicit expression of the output state produced at the \( k \)-th round can be obtained as follows. We regard the recursive protocol as a quantum instrument, with outcomes in the set \( \{1, \ldots, K+1\} \) corresponding to pure quantum operations with Kraus operators

\[
M_k := \begin{cases} B^{(k)}_{\text{suc}} B^{(k-1)}_{\text{fail}} \cdots B^{(1)}_{\text{fail}} & k = 1, \ldots, K, \\ B^{(K)}_{\text{fail}} B^{(K-1)}_{\text{fail}} \cdots B^{(1)}_{\text{fail}} & k = K + 1 \end{cases} \quad (36)
\]

For \( k \leq K \), the Kraus operators are characterized in Appendix B, where we prove the following relation

\[
M_k |\varphi_E\rangle = \begin{cases} 0 & E \in U_{k-1} \\ \sqrt{(r_k - r_{k-1}) \frac{q_E}{p_E}} |\psi_E\rangle & E \in \text{Sp}(\varphi) \setminus U_{k-1} \end{cases}, \quad (37)
\]
Having set $r_0 := 0$. The expressions of the fidelity and of the probability of success, anticipated in Eqs. (32) and (33), can be easily derived from the above equation, using the relations $\rho_{\text{succ}}^{(k)} = \|M_k|\varphi\|^2$ and $F_{\max}^{(k)} = |\langle \psi | M_k | \varphi \rangle|^2 / \|M_k | \varphi\|^2$.

The output state in case of success at the $k$-th round is given by

$$|\psi^{(k)}\rangle := \frac{M_k |\varphi\rangle}{\|M_k |\varphi\|} = \frac{1}{\sqrt{F_{\max}^{(k)}}} \bigoplus_{E \in \text{Sp}(\varphi) \setminus U_{k-1}} \sqrt{q_E} |\psi_E\rangle. \tag{38}$$

Note that $|\psi^{(k)}\rangle$ is a truncated version of the target state, in which the energy spectrum has been deprived of all the values in $U_{k-1}$ and of all the values that are not in in the spectrum of $|\varphi\rangle$. Essentially, 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 $L$. Protocols with more than $L$ rounds terminate after $L$ steps, meaning that the probability of success satisfies

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

The fact that the protocol is guaranteed to terminate in a finite time is an appealing feature, which is generic to all quantum loop programs [60]. It is worth stressing that the termination time $L$ 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. Consider for example the case of $N$ identical non-interacting 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

$$L \leq \left( d + \frac{N - 1}{N} \right) < (N + 1)^{d-1},$$

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], the recursive protocol is guaranteed to reach unit probability in a polynomial number of iterations.

D. Coherent coarse-graining of the recursive protocol

Given a quantum filter, coherent coarse-graining allows us to increase the fidelity of the favourable subset of quantum operations while keeping the same probability of success. Let us apply this method to the instrument resulting from our $K$-round recursive protocol. By coherently coarse-graining the quantum operations corresponding to the first $T$ rounds, we obtain a binary pure filter with successful Kraus operator

$$M^{(T)} := U \sum_{k=1}^{T} M_k^\dagger M_k \tag{39}$$

where $U$ is a unitary that commutes with the energy and satisfies the relation $U |\varphi_E\rangle = |\psi_E\rangle$ for every $E$ [cf. Eq. (24)]. For $T \leq L$, we then obtain

$$M^{(T)} |\varphi_E\rangle = \begin{cases} |\psi_E\rangle & E \in U_T \\ \sqrt{r_T} \frac{q_E}{p_E} |\psi_E\rangle & E \notin \text{Sp}(\varphi) \setminus U_T, \end{cases} \tag{40}$$

having used Eq. (37) for the Kraus operators of the original protocol. Note that, as long as $T \leq L$, the coherently coarse-grained Kraus operators depend only on $T$, and not on the number on the number of rounds in the original protocol. For $T = L$, coherent coarse-graining yields a deterministic process: precisely, it yields the optimal energy-preserving channel for the transformation $|\varphi\rangle \to |\psi\rangle$, as one can see from corollary 1).

Let us evaluate the performances of the coherently coarse-grained protocol. By construction, the probability of success is equal to the probability $p_{\text{succ}}(T)$ given in Eq. (33). On the other hand, the fidelity can be evaluated explicitly from Eq. (40), which yields

$$F'(T) = \left[ \frac{\sum_{E \in U_T} \sqrt{p_E q_E} + \sqrt{r_T} \sum_{E \in \text{Sp}(\varphi) \setminus U_T} q_E}{\sum_{E \in U_T} p_E + r_T \sum_{E \in \text{Sp}(\varphi) \setminus U_T} q_E} \right]^2. \tag{41}$$

From the proof of theorem 2 one can see that the fidelity of the coherent coarse-graining is strictly larger than the fidelity of the recursive protocol. 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 provide apply the recursive protocol and the method of coherent coarse-graining to the tasks of parameter estimation, cloning of quantum clocks, phase-insensitive amplification of coherent states, and approximate correction in ancilla-driven quantum computation.
A. Phase estimation

In the following we apply the recursive protocol to the task of phase estimation \cite{40,55}. Consider a pure quantum state of the form $|\varphi\rangle = e^{-i\theta H}|\varphi\rangle$, where $H = \sum_{n=0}^{N-1} n|n\rangle\langle n|$ is the Hamiltonian of the system, $|\varphi\rangle$ is a fixed initial state, and $\theta$ is an angle between 0 and $2\pi$. The task here is to estimate the unknown parameter $\theta$ by performing a suitable quantum measurement. Consider the maximally coherent state \cite{14}

$$|\varphi\rangle := \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} |n\rangle,$$  \hfill (42)

consisting of a uniform superposition of the $N$ energy eigenstates. To evaluate the precision of estimate $\theta$, we use the gain function $G(\theta, \hat{\theta})$ defined by \cite{40}

$$G(\theta, \hat{\theta}) := \frac{1 + \cos(\theta - \hat{\theta})}{2}.$$  \hfill (43)

The task of optimal phase estimation is to maximize the average gain

$$\langle G \rangle := \int \frac{d\theta}{2\pi} \int \frac{d\hat{\theta}}{2\pi} G(\theta, \hat{\theta}) \langle \varphi_\theta | E_\hat{\theta} | \varphi_\theta \rangle$$  \hfill (44)

over all possible POVMs $\{E_{\hat{\theta}}\}$. For phase estimation with pure states, the optimal POVM has been derived by Holevo \cite{40}. If the probe $|\varphi_\theta\rangle$ is measured without passing through the filter, the optimal POVM gives the gain (43)

$$\langle G \rangle = \frac{1}{2} + \frac{1}{2} \langle \varphi_\theta | \Delta | \varphi_\theta \rangle$$

$$\Delta_{ij} = \frac{1}{2} \left[ \delta_{i(j-1)} + \delta_{i(j+1)} \right].$$  \hfill (45)

We referred to this quantity as $\langle G_{opt} \rangle$, the gain for deterministic estimation. By direct calculation, one finds that $\langle G_{det} \rangle$ has the standard quantum limit scaling

$$\langle G_{det} \rangle = 1 - \frac{1}{2N}.$$  \hfill (46)

Now, the precision of phase estimation depends highly on the probe state. The state that maximizes the gain is the “sine-state” \cite{62}, given by

$$|\varphi_{opt}\rangle = \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} \sin\left(\frac{n\pi}{N}\right) |n\rangle.$$  \hfill (47)

In the probabilistic case one can try to transform the state $|\varphi_\theta\rangle$ into the state $|\varphi_{opt,\theta}\rangle := e^{i\theta H}|\varphi_{opt}\rangle$. Since the energy-preserving operations are covariant, implementing the transformation $|\varphi_\theta\rangle \rightarrow |\varphi_{opt,\theta}\rangle$ for an unknown $\theta$ is equivalent to implementing the transformation $|\varphi\rangle \rightarrow |\varphi_{opt}\rangle$.

So far, all studies about probabilistic phase estimation have focussed on binary filters \cite{24-27}. Here we apply our recursive protocol to approximate the desired conversion $|\varphi_\theta\rangle \rightarrow |\varphi_{opt,\theta}\rangle$, using the output to perform phase estimation. The key point here is that, conditional on the outcomes of the filters, the output states of our protocol are pure at every step. Thanks to this fact, we can apply Holevo’s recipe for optimal phase estimation \cite{40}. Denoting by $M_k$ the Kraus operator corresponding to success at the $k$-th round, we have that the gain is given by

$$\langle G^{(k)} \rangle = \frac{1}{2} + \frac{1}{2p_{\text{succ}}} \langle \varphi_{opt} | M_k^\dagger M_k | \varphi_\theta \rangle$$  \hfill (48)

where $p_{\text{succ}}$ is the probability of success for the $k$-th round, given by Eq. (33). Averaging over all rounds, we obtain the gain

$$\langle G_K \rangle := \frac{\sum_{k=1}^{K} p_{\text{succ}}^{(k)} \langle G^{(k)} \rangle}{p_{\text{succ}}(K)}$$  \hfill (49)

where $p_{\text{succ}}(K)$ is the probability of achieving success in one of the $K$ rounds. Combining Eqs. (47) and (48), the average gain can be expressed as

$$\langle G_K \rangle = 1 - \pi^2 \left[ K(K-1) + \frac{1}{4} \right] + O\left( \frac{K}{N} \right)^3.$$  \hfill (50)

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 $K$. The success probability can also be evaluated analytically in the regime $N \gg K$ and its value is given by

$$p_{\text{succ}}(K) = \frac{1}{2} + \frac{\pi^2}{N^2} \left[ K(K-1) + \frac{1}{8} \right] + O\left( \frac{K}{N} \right)^3.$$  \hfill (51)

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

$$\langle G_K \rangle = 1 - \frac{3\pi^2}{16N^2} - \frac{\delta(K)}{2}$$

$$\delta(K) := p_{\text{succ}}(K) - 1/2.$$  \hfill (52)

In addition to the gain of the $K$-round recursive protocol, we evaluate the gain of its coherent coarse-graining. By
coherently coarse-graining over the first $T$ rounds, we obtain the average gain given by

$$
\langle G'_T \rangle = \frac{1}{2} + \frac{1}{2p_{\text{succ}}(T)} \langle \varphi_0 | M^{(T)} | \Delta M^{(T)} | \varphi_0 \rangle ,
$$

with $M^{(T)}$ as in Eq. (39). In the $N \gg K$ regime, the gain has the analytical expression

$$
\langle G' \rangle = 1 - \frac{\pi^2}{4N^2} \{1 + 4[\delta(K)]^2\} ,
$$

with $\delta$ defined as in Eq. (52).

It is interesting to compare the performance of the coherent coarse-graining with the optimal tradeoff 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 tradeoff. The answer turns out to be negative: by evaluating Eq. (53) 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. (45), clearly indicating sub-optimality.

![FIG. 2. Energy-preserving estimation via the recursive protocol and its coherent coarse-graining.](image)

The figure shows the tradeoff between success probability and average gain for phase estimation with the maximally coherent state $|\varphi_0\rangle := 1/\sqrt{N} \sum_{n=0}^{N-1} e^{i\phi_n} |n\rangle$ for $N = 61$. The green solid line (with numerics represented by red dots) shows the probability-gain tradeoff for a recursive protocol with $K = 17$ rounds. At the first round the protocol reaches the maximum possible gain, equal to $G_{\max} = 99.9\%$, in agreement with the analytical expression of Eq. (50). The blue solid line (with numerics represented by the black dots) shows the tradeoff for filters generated by coherent coarse-graining, with the $T$-th point corresponding to the coherent coarse-graining of the first $T$ quantum operations of the recursive protocol. Notice that the coherent coarse-graining remains approximately equal to $G_{\max} = 99.9\%$. Finally, the black dashed line represents the gain for the optimal deterministic estimation protocol.

In Figure 2 we illustrate the tradeoff between the probability of success and the average gain for a 17-rounds recursive protocol and for its coherent coarse-graining. It can be seen that the recursive protocol manages to increase the probability of success by approximately 30% from the first round to the 14-th, while keeping the average gain above the deterministic gain. In Figure 3 we show the scaling of the gain and the success probability with the dimension $N$ for recursive protocols with different numbers of rounds.

![FIG. 3. The gain and success probability curves](image)

In the above figures 3(a) the growth of average gain $G$ and 3(b) the decrease of the total success probability as a function of the dimension $N$ are plotted. Here different lines represent different numbers of filtering rounds performed, correspondingly $K = 1$ (black line with black dots), $K = 2$ (green line with red dots) and $K = 3$ (blue line with purple dots). In figure 3(a) the dashed line represents the unit gain upper bound.

### B. Energy-preserving cloning of quantum clocks

In quantum cloning [63, 64], one is given $N$ identical copies of an unknown quantum state and wants to produce a larger number $M \geq N$ of approximate copies. For a given set of states, the key question is how the accuracy of the copies depends on $N$ and $M$. Typically, this question has been addressed without imposing any constraint on the cloning process, except for its compatibility with the laws of quantum mechanics. However, in sev-
eral situations one may be constrained to use processes that preserve the energy. Consider for example a scenario where one wants to clone the states 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 while preserving the energy, 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 conversion. 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 $\langle M \psi|H|\psi\rangle$, meaning that the black state should have energy close to $\langle (M - N)\psi|H|\psi\rangle$.

The problem of energy-preserving cloning is equivalent to the conversion problem considered earlier in this paper: 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 $|\psi\rangle^{\otimes M}$. Since every energy-preserving operation is covariant under time evolution, maximizing the fidelity for the conversion $|\psi\rangle^{\otimes N} \otimes |\beta\rangle \rightarrow |\psi\rangle^{\otimes M}$ 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: we consider a two-level quantum clock starting at time $t = 0$ in the coherent superposition $|\psi\rangle = |\alpha\rangle^{\otimes N} \otimes |\beta\rangle$, and subsequently evolving with the time independent Hamiltonian $H = \hbar \sigma_x/2$, where $\sigma_x = |+\rangle\langle+|-\rangle\langle-|$. Note that the average energy of the clock remains zero during the evolution, i.e. $\langle \psi|H|\psi\rangle = 0$ for every time $t$.

Let us focus on the task of $N$-to-$M$ cloning in the case where the difference $M - N$ is even. Under this condition, we can choose the blank state to be an energy eigenstate with energy exactly equal to zero. Specifically, we choose the state $|\beta\rangle = |(M - N)/2, 0\rangle$, belonging to the symmetric eigenbasis

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

where $S_L$ denotes the set of all permutations of a $L$-element set and $U_\pi$ is the unitary that permutes $L$ Hilbert spaces according to the permutation $\pi$. We now apply our recursive protocol, in which the input state is filtered step by step, producing a sequence of output states approximating $M$ ideal copies. We expand the states $|\psi\rangle^{\otimes N}$ and $|\psi\rangle^{\otimes M}$ as

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

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

$$F^{(1)}_{\text{max}} = \frac{1}{2^M} \sum_{n = -N}^N \left( \frac{M}{M-n} \right).$$

The above fidelity turns out to be equal to the absolute maximum of the fidelity achievable over all covariant quantum operations, derived by Fiurašek in Ref. [28]. 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 tradeoff curves between the fidelity and the probability of success are plotted in Figure 4 for the case of $N = 80, M = 400$ and $K = 32$, using Eqs. (32) and (33). In addition, we compare the fidelity of the recursive protocol with that of its coherent coarse-graining, given by Eq. (41). 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 Figure 4). The deterministic fidelity, derived from corollary 1, coincides with the fidelity for phase-covariant cloning [34], meaning that the optimal cloner can be realized in an energy-preserving fashion.

Figure 4 illustrates very well the advantages of the recursive protocol. At the first round the fidelity is very high, but the success probability has the minuscule value $p_{\text{success}}^{(1)} = 6 \times 10^{-20}$. The subsequent rounds of the protocol increase the success probability dramatically, reaching a probability of approximately 23% at the 31-th step. Quite interestingly, the fidelity remains above the maximal deterministic fidelity up to almost the very last step.

C. Phase insensitive 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 non-degenerate and the number-preserving quantum operations have diagonal Kraus operators in the Fock basis $\{|n\}\). In the following we consider the application of the recursive protocol to the amplification of the coherent state of light

$$|r_1\rangle \rightarrow |r_2\rangle, \quad 0 \leq r_1 \leq r_2.$$

Note that, since we require the amplification map to be part of a number-preserving quantum instrument, our
the obstacle, we define a threshold protocol cannot be applied directly. To overcome the tradeoff with the photon number modulation.

The recursive protocol maintains fidelity larger than the optimal deterministic cloning protocol. Notice that the recursive protocol holds fidelity larger than the optimal deterministic cloning protocol.

The fact that the number is preserved only on average gives one the opportunity to reach high fidelity in a probabilistic fashion. In the case of amplifiers, the tradeoff of the number seems to be a daunting task. However, the recursive protocol and its coherent coarse-graining give fidelities greater than the optimal deterministic cloning protocol. Notice that the difference between the two curves becomes large as the success probability tends to one. 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.

of success in Eq. (33) can be expressed as

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

Interestingly, the successful quantum operation at the first round of our protocol ($k = 1$) coincides with the optimal probabilistic amplifier for coherent states on a circle [28, 31], which indeed can be implemented with energy-preserving operations. Specifically, evaluating Eqs. (54) and (55) for $k = 1$ one retrieves the expressions for the optimal fidelity and success probability appearing in Eqs. (6.24) and (6.36) of Ref. [31]. 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 Figure 5 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 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.
\[ F = 83.9\%. \] Note also the very high fidelities of amplification that are obtained despite the number-preserving constraint, at the price of very small success probabilities. 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 \( F_{\text{universal}} = 4/9 \) \cite{30, 65}.

\section*{D. Energy-preserving correction in ancilla-driven quantum computation}

In ancilla-driven quantum computation \cite{33} 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 \cite{66}. However, in many non-ideal situations the measurement extracts some information, thus inducing a non-unitary evolution on the system. When this is the case, one can attempt to correct the unwanted non-unitarity by performing additional measurements. This type of correction has been studied in Refs. \cite{67, 68}, 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 \( d \) non-degenerate energy levels interacts with an ancilla via an energy-preserving unitary evolution. Then, the ancilla undergoes the measurement of an observable that is compatible with the energy. As a result, the system evolves randomly according to an energy-preserving instrument \( \{ M_x \}_{x \in X} \). We assume that the measurement on the ancilla is a rank-one projective measurement and, therefore, the quantum operations \( \{ M_x \} \) are pure. For every given \( x \in X \), the problem is to correct the quantum operation \( M_x \), 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 \( \{ N^{(x)}_{\text{succ}}, N^{(x)}_{\text{fail}} \} \). Due to the presence of the filter, an initial pure state \( |\eta\rangle \) is transformed probabilistically into the pure state

\[ |\eta_x\rangle = \frac{N^{(x)}_{\text{succ}} M_x |\eta\rangle}{\| N^{(x)}_{\text{succ}} M_x |\eta\rangle \|} . \]

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

\[ p(\eta|x, \text{succ}) \, d\eta = \lambda_x \left\| N^{(x)}_{\text{succ}} M_x |\eta\rangle \right\|^2 \, d\eta , \]

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

\[ F_x := \int d\eta \, p(\eta|x, \text{succ}) \left| \langle \eta_x | U_x |\eta\rangle \right|^2 \]

\[ = \frac{\int d\eta \left| \langle \eta | U_x^\dagger N^{(x)}_{\text{succ}} M_x |\eta\rangle \right|^2}{\int d\eta' \| N^{(x)}_{\text{succ}} M_x |\eta'\rangle \|^2} \]

\[ = \frac{F^{(x)}_0 \cdot d + 1}{d + 1} \] \hspace{1cm} (56)

where \( F^{(x)}_0 \) is the fidelity given by

\[ F^{(x)}_0 = \left\| \left( e_0 | U_x^\dagger N^{(x)}_{\text{succ}} M_x | e_0 \right) \right\|^2 \]

\[ = \frac{\sum_{n=1}^d \| e_0 \|^2 \| M_x | e_0 \rangle \|^2}{\| N^{(x)}_{\text{succ}} M_x | e_0 \rangle \|^2} . \]

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

\[ |\varphi_x\rangle := \frac{M_x | e_0 \rangle}{\| M_x | e_0 \rangle \|} \rightarrow |\psi_x\rangle := U_x | e_0 \rangle . \]

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 an efficient 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.

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

\[ F^{(k)}_x = \frac{d + 2 - k}{d + 1} , \]

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

\[ p^{(k)}_{\text{succ}} = \left\{ \begin{array}{ll} \mu^{d-k}(1 - \mu)^2(d + 1 - k)/(1 - \mu^d) & \quad k > 1 \\ \mu^{d-1}(1 - \mu)d/(1 - \mu^d) & \quad k = 1 . \end{array} \right. \]

The features of the recursive protocol and of its coherent coarse-graining are illustrated in figure 6 for \( d = 100, \mu = 0.9 \) and \( K = 70 \). The probability of success increases from a very small value \( p^{(1)}_{\text{succ}} = 3 \times 10^{-4} \) to approximately 14% at the 68-th step, at the cost of a reduced fidelity.
In this paper we investigated how quantum coherence across different energy levels can be manipulated without exchanging energy between the system and the environment. Every interaction satisfying this restriction can be engineered in such a way that the environment stays in the same energy eigenspace for the whole time evolution. In addition to the deterministic evolutions, we considered probabilistic evolutions driven by measurements. In this scenario the system evolves stochastically through a set of quantum maps, which preserve the energy on average but individually are able to modulate the energy distribution of the system. Exploiting such a modulation, one can achieve transformations that are impossible deterministically.

Within the energy-preserving paradigm, we asked how well a given input state can be transformed into a desired target and what is the maximum probability that such a transformation occurs. We first addressed the search of the optimal quantum operations with a fixed value of the success probability, showing that pure quantum operations are always optimal for this task. In particular, we observed that unitary transformations are always optimal among the set of deterministic transformations. The characterization of the optimal quantum operations allowed us to construct a multi-round recursive protocol that achieves maximum fidelity with maximum success probability in each round. For a system with finite energy spectrum, the protocol terminates with unit probability in a finite number of steps. In particular, the protocol terminates in less than $O(N^{d-1})$ rounds for a system of $N$ non-interacting particles with $d$ energy levels.

In addition to the recursive protocol, we introduced the operation of coherent coarse-graining, by which a set of coexisting quantum operations are joined into a single, pure quantum operation, with the same probability of occurrence of the original set and, typically, with a higher fidelity with the target. By coherently coarse-graining the quantum operations of our recursive protocol we showed that one can efficiently approximate the trade-off curve between fidelity and success probability. Remarkably, when applied to the problem of phase estimation with maximally coherent states, this method yields points that lie exactly on the optimal trade-off curve, provided that the number of energy levels is sufficiently large. We conjecture that the coherent coarse-graining of our recursive filter yields the optimal tradeoff curve in all asymptotic scenarios where the energy distribution of the state is sufficiently regular, including in particular the cases of phase estimation with multiple copies and quantum cloning in the regime where the number of copies is asymptotically large [69].

VI. CONCLUSIONS

In this paper we focused on the optimal transformations of pure states. However, for many practical applications it is important to consider transformations of mixed states. While the general problem is hard to treat analytically, our result provides tools to study on how an input mixed state can be converted into (an approximation of) a target pure state under the energy-preserving restriction. This version of the problem is well motivated—possibly even more motivated than the task of transforming a mixed state into another mixed state—and has applications in tasks like mixed state purification [70], super-broadcasting [71], and in the evaluation of linear quantum benchmarks [72]. In addition to the generalization to mixed states, another 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 study 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 [48, 51, 73–75]. While the characterization of the best angular-momentum-preserving transformations is largely open, our result suggests a possible strategy to approach the problem, by considering probabilistic modulation of the amplitudes of the wave-function in sectors with different angular momentum. Also in this case, one can construct first the adaptive protocol and then 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 this work can open the way to a systematic optimization of quantum operations under arbitrary conservation laws.
ACKNOWLEDGMENTS

This work is supported by the National Basic Research Program of China (973) 2011CBA00300 (2011CBA00301), by the National Natural Science Foundation of China (11450110096, 61033001, 6136136003), by the Foundational Questions Institute (FQXi-RFP3-1325), and by the 1000 Youth Fellowship Program of China.

[1] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. 96, 010401 (2006), URL http://link.aps.org/doi/10.1103/PhysRevLett.96.010401.

[2] V. Giovannetti, S. Lloyd, and L. Maccone, Science 306, 1330 (2004), URL http://www.sciencemag.org/content/306/5700/1330.abstract.

[3] V. Giovannetti, S. Lloyd, and L. Maccone, Nat Photon 5, 222 (2011), URL http://dx.doi.org/10.1038/nphoton.2011.35.

[4] W. G. Unruh, Phys. Rev. A 51, 992 (1995), URL http://link.aps.org/doi/10.1103/PhysRevA.51.992.

[5] S. J. Van Enk and H. J. Kimble, Quantum Info. Comput. 2, 1 (2002), ISSN 1533-7146, URL http://dx.doi.org/10.1006/qmci.2001.0117.

[6] J. Gea-Banacloche, Phys. Rev. A 65, 022308 (2002), URL http://link.aps.org/doi/10.1103/PhysRevA.65.022308.

[7] J. Gea-Banacloche and M. Ozawa, Phys. Rev. A 74, 060301 (2006), URL http://link.aps.org/doi/10.1103/PhysRevA.74.060301.

[8] S. Huelga and M. Plenio, Contemporary Physics 54, 181 (2013), URL http://dx.doi.org/10.1080/00107514.2013.829687.

[9] M. B. Plenio, E. Demler, and G. Vidal, Phys. Rev. Lett. 96, 010401 (2006), URL http://link.aps.org/doi/10.1103/PhysRevLett.96.010401.

[10] V. Giovannetti, S. Lloyd, and L. Maccone, Science 306, 1330 (2004), URL http://www.sciencemag.org/content/306/5700/1330.abstract.

[11] V. Giovannetti, S. Lloyd, and L. Maccone, Nat Photon 5, 222 (2011), URL http://dx.doi.org/10.1038/nphoton.2011.35.

[12] J. Gea-Banacloche, Phys. Rev. A 65, 022308 (2002), URL http://link.aps.org/doi/10.1103/PhysRevA.65.022308.

[13] J. Gea-Banacloche and M. Ozawa, Phys. Rev. A 74, 060301 (2006), URL http://link.aps.org/doi/10.1103/PhysRevA.74.060301.

[14] S. Huelga and M. Plenio, Contemporary Physics 54, 181 (2013), URL http://dx.doi.org/10.1080/00107514.2013.829687.

[15] M. B. Plenio, E. Demler, and G. Vidal, Phys. Rev. Lett. 96, 010401 (2006), URL http://link.aps.org/doi/10.1103/PhysRevLett.96.010401.

[16] V. Giovannetti, S. Lloyd, and L. Maccone, Science 306, 1330 (2004), URL http://www.sciencemag.org/content/306/5700/1330.abstract.

[17] V. Giovannetti, S. Lloyd, and L. Maccone, Nat Photon 5, 222 (2011), URL http://dx.doi.org/10.1038/nphoton.2011.35.
Appendix A: Derivation of the optimal protocol

The optimal filter in the $k$-th round is determined by induction from the requirements 1-3 in subsection IV A. At the first round, the filter attempts at converting $|\varphi\rangle$ into $|\psi\rangle$. According to corollary 2, the maximum fidelity is given by

$$F^{(1)}_{\text{max}} = \sum_{E \in \text{Sp}(\varphi)} q_E$$

and can be achieved with probability

$$P_{\text{succ}}^{(1)} = \left( \min_{E \in \text{Sp}(\varphi) \cap \text{Sp}(\psi)} \frac{P_E}{q_E} \right) F^{(1)}_{\text{max}}.$$

The optimal quantum operation must be pure and that its Kraus operator $B^{(1)}_{\text{succ}}$ must satisfy the condition

$$B^{(1)}_{\text{succ}} |\varphi_E\rangle = |\psi_E\rangle \quad \forall E \in R_1,$$

where $R_1$ is the set of all energy values in $\text{Sp}(\varphi) \cap \text{Sp}(\psi)$ that minimize the ratio $P_E/q_E$ [cf. Eq. (25)]. Writing the unsuccessful quantum operation in the Kraus form

$$B^{(1)}_{\text{fail}}(\cdot) = \sum_{t} B^{(1)}_{\text{fail},t} \cdot B^{(1)}_{\text{fail},t}^\dagger,$$

we then obtain

$$B^{(1)}_{\text{fail},t} |\varphi_E\rangle = 0 \quad \forall E \in R_1$$

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

$$\rho^{(2)} = \sum_{t} P_t^{(2)} \left| \varphi^{(2),t}\right \rangle \left\langle \varphi^{(2),t} \right|.$$
with \( p_t^{(2)} := \|B_{\text{fail},t}^{(1)}\varphi\|^2 / \sum_{t'} \|B_{\text{fail},t'}\varphi\|^2 \) and
\[
\varphi^{(2,t)} := \frac{B_{\text{fail},t}\varphi}{\|B_{\text{fail},t}\varphi\|}. \tag{A3}
\]

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
\[
|\varphi^{(2,t)}\rangle = \bigoplus_E \sqrt{p_E^{(2,t)}} |\varphi_E^{(2,t)}\rangle,
\]
for suitable probabilities \( \{p_E^{(2,t)}\} \) and suitable energy eigenstates \( \{|\varphi_E^{(2,t)}\rangle\} \). Note that, due to the condition in Eq. (A2), one has
\[
S_p(\varphi^{(2,k)}) \subseteq S_p(\varphi) \setminus R_1. \tag{A4}
\]

Using this fact, we can upper bound the fidelity achievable from the state \( |\varphi^{(2,t)}\rangle \)—call it \( F_{\text{max}}^{(2,t)} \)—as
\[
F_{\text{max}}^{(2,t)} \leq \sum_{E \in S_p(\varphi^{(2,t)})} q_E
\]
\[
\leq \sum_{E \in S_p(\varphi) \setminus R_1} q_E
\]
the first inequality coming from Corollary 2. In turn, this allows us to upper bound the overall fidelity at the second step as
\[
F^{(2)} \leq \max_t F_{\text{max}}^{(2,t)}
\]
\[
\leq \sum_{E \in S_p(\varphi) \setminus R_1} q_E
\]
\[
=: F_{\text{max}}^{(2)} \tag{A5}
\]

The bound is attained when the quantum operation \( B_{\text{fail}}^{(1)} \) is pure and its Kraus operator is given by
\[
B_{\text{fail}}^{(1)} := \sqrt{I - B_{\text{suc}}^{(1)\dagger}B_{\text{suc}}^{(1)}}. \tag{A6}
\]

Luckily, this choice maximizes not only the fidelity at the second step, but also the probability that maximum fidelity is achieved: indeed, denoting by \( p_{\text{suc}}^{(2)} \) the probability that the output has fidelity \( F_{\text{max}}^{(2)} \) with the target and by \( p_{\text{suc}}^{(2,t)} \) the probability that the optimal transformation takes place on the state \( |\varphi^{(2,t)}\rangle \), we have the bound
\[
p_{\text{suc}}^{(2)} \leq \sum_t p_{\text{suc}}^{(2,t)} p_{\text{suc}}^{(2,t)}
\]
\[
= \sum_t p_{\text{suc}}^{(2)} \left[ \min_{E \in S_p(\varphi^{(2,t)}) \setminus S_p(\psi)} \frac{p_E^{(2,t)}}{q_E} \right] F_{\text{max}}^{(2,t)}
\]
\[
\leq \min_E \left[ \sum_t p_{\text{suc}}^{(2)} p_E^{(2,t)} \frac{F_{\text{max}}^{(2)}}{q_E} \right] F_{\text{max}}^{(2)} \tag{A7}
\]
the equality in the second line coming from corollary 2. It is easy to verify that the pure quantum operation of Eq. (A6) reaches the bound: indeed, its output state
\[
|\varphi^{(2)}\rangle = \frac{B_{\text{fail}}^{(1)}\varphi}{\|B_{\text{fail}}^{(1)}\varphi\|},
\]
can be converted optimally into the state \( |\psi\rangle \) with probability given by corollary 2, which now yields
\[
p_{\text{suc}}^{(2)} = \min_{E \in S_p(\varphi^{(2)}) \setminus S_p(\psi)} \left[ \frac{p_E^{(2)}}{q_E} \right] F_{\text{max}}^{(2)} \tag{A8}
\]
with \( p_E^{(2)} := \|P_E|\varphi^{(2)}\|^2 \). Inserting this equality in Eq. (A8) we then obtain that the bound of Eq. (A7) 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 \( B_{\text{suc}}^{(k)} \) and \( B_{\text{fail}}^{(k)} \), respectively. Expanding the state at the \( k \)-th step as
\[
|\varphi^{(k)}\rangle = \bigoplus_E \sqrt{p_E^{(k)}} |\varphi_E^{(k)}\rangle,
\]
the successful Kraus operator is determined in an essentially unique way by corollary 2, which yields the condition
\[
B_{\text{suc}}^{(k)} |\varphi_E^{(k)}\rangle = \left[ \min_{E' \in S_p(\varphi^{(k)}) \setminus S_p(\psi)} \sqrt{\frac{p_{E'}}{q_{E'}}} \right] \sqrt{\frac{q_E}{p_E}} |\psi\rangle \tag{A9}
\]
for every energy \( E \) in \( S_p(\varphi^{(k)}) \). The unsuccessful Kraus operator is then given by
\[
B_{\text{fail}}^{(k)} := \sqrt{I - B_{\text{suc}}^{(k)\dagger}B_{\text{suc}}^{(k)}}, \tag{A10}
\]
and its definition if essentially unique, up to the application of an energy-preserving unitary on the output and to a possible re-definition of \( B_{\text{fail}}^{(k)} \) outside the relevant subspace.

Applying iteratively Eqs. (A9) and (A10) it is easy to obtain that the eigenstates \( |\varphi_E^{(k)}\rangle \) are independent of \( k \), i.e. one has
\[
|\varphi_E^{(k)}\rangle \equiv |\varphi_E\rangle \quad \forall k = 1, \ldots, K, \forall E \in S_p(\varphi^{(k)}).
\]
This condition implies that Eq. (A9) can be rewritten as
\[
B_{\text{suc}}^{(k)} |\varphi_E\rangle = \left[ \min_{E' \in S_p(\varphi^{(k)}) \setminus S_p(\psi)} \sqrt{\frac{p_{E'}}{q_{E'}}} \right] \sqrt{\frac{q_E}{p_E}} |\psi\rangle
\]
for every energy \( E \) in \( S_p(\varphi^{(k)}) \).
Appendix B: Kraus operators of the recursive protocol

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

$$M_k = B^\text{succ}_k B^\text{fail}_{k-1} \cdots B^\text{fail}_1. \quad (B1)$$

To characterize $M_k$ we first analyze the operators $B^\text{fail}_i$, with $i = 1, \ldots, k-1$. Combining Eqs. (A9) and (A10), we obtain that $B^\text{fail}_i$ satisfies the condition

$$B^\text{fail}_i |\varphi_E\rangle = 0 \quad \forall E \in R^\text{fail}_i, \quad (B2)$$

where $R^\text{fail}_i$ is the set defined by

$$R^\text{fail}_i := \left\{ E \in \text{Sp}(\varphi(i)) \cap \text{Sp}(\psi) \mid \frac{p^\text{fail}_i}{q_E} = r^\text{fail}_i \right\}, \quad (B3)$$

$r^\text{fail}_i$ being the minimum non-zero value of the ratio $r^\text{fail}_i = p^\text{fail}_i/q_E$. Now, the key observation is provided by the following

**Lemma 2.** The set $R^\text{fail}_1$ coincides with the set $R_i$ defined in Eq. (30).

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

$$p^\text{ fail}_i(q_E) = \frac{\| B^\text{fail}_i |\varphi(i)\rangle \|^2}{q_E \| B^\text{fail}_i |\varphi(i)\rangle \|^2} = \frac{\langle \varphi(i) | P_E \left( I - B^\text{succ}_i B^\text{suc}_i \right) P_E |\varphi(i)\rangle}{q_E \| B^\text{fail}_i |\varphi(i)\rangle \|^2} \quad (B4)$$

where the first equality follows from the definition $|\varphi(i+1)\rangle := B^\text{fail}_i |\varphi(i)\rangle / \| B^\text{fail}_i |\varphi(i)\rangle \|$, the second from the relation $B^\text{ fail}_i := \sqrt{I - B^\text{succ}_i B^\text{suc}_i}$, and the third from Eq. (A9). Eq. (B4) shows that the set of energies for which the ratio $p^\text{ fail}_i(q_E)$ has the smallest value coincides with the set of energies for which the ratio $p^\text{ fail}_i(q_E)$ has the second smallest value, which in turn coincides with the set of energies for which the ratio $p^\text{ fail}_i(q_E)$ has the $(i+1)$-th smallest value. By definition of the sets $R^\text{ fail}_1$ and $R_i$, this proves the thesis.

Using the above lemma, Eq. (B2) becomes

$$B^\text{fail}_i |\varphi_E\rangle = 0 \quad \forall E \in R_i. \quad (B5)$$

We now use this relation to determine the form of the Kraus operator $M_k$. Setting $m^k_E := \| M_k |\varphi_E\rangle \|^2$, the definition of $M_k$ [Eq. (B1)] gives the bound

$$m^k_E \leq \| B^\text{fail}_i |\varphi_E\rangle \|^2 \quad \forall i = 1, \ldots, k-1.$$

Combining this bound with Eq. (B5) we obtain the condition

$$m^k_E = c_k q_E \quad \forall E \in \text{Sp}(\varphi) \setminus U_{k-1}, \quad (B6)$$

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. (B1), the last operator acting in the sequence is $B^\text{suc}_k$. Hence, we know that the initial amplitude $\sqrt{q_E}$ should be modulated to $\sqrt{q_E}$ for all the energy eigenvalues that survived the first $k-1$ rounds, that is,

$$m^k_E = c_k q_E \quad \forall E \in \text{Sp}(\varphi) \setminus U_{k-1}. \quad (B7)$$

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

$$\sum_{k=1}^{K+1} m^k_E = 1 \quad \forall E.$$

Combining this fact with Eqs. (B6) and (B7) we then obtain the recursion relations

$$c_1 q_E \equiv 1 \quad \forall E \in R_1$$

$$(c_1 + c_2) \frac{q_{E}}{p_{E}} = 1 \quad \forall E \in R_2$$

$$\vdots$$

$$(\sum_{k=1}^{K} c_k) \frac{q_{E}}{p_{E}} = 1 \quad \forall E \in R_K.$$

Finally, using the definition of the sets $R_k$ [Eq. (30)] 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

$$M_k |\varphi_E\rangle = \begin{cases} 0 & E \in U_{k-1} \\
\sqrt{(r_k - r_{k-1}) \frac{q_{E}}{p_{E}}} |\psi_E\rangle & E \in \text{Sp}(\varphi) \setminus U_{k-1}, \end{cases} \quad (B8)$$

this proving the expression in Eq. (37).