

Quantum observables over time for information recovery

Gabriele Bressanini ^{1,*} Farhan Hanif,¹ Hyukjoon Kwon ² and M. S. Kim^{1,2}

¹*Blackett Laboratory, Imperial College London, London SW7 2AZ, United Kingdom*

²*Korea Institute for Advanced Study, Seoul 02455, South Korea*



(Received 19 December 2024; accepted 18 May 2026; published 19 August 2026)

We introduce the concept of quantum observables over time (QOOT), an operator that jointly describes two observables at two distinct time points, as a dual of the quantum state over time formalism. We provide a full characterization of the conditions under which a QOOT can be properly defined, via a no-go theorem. We use QOOTs to establish a notion of time reversal for generic quantum channels with respect to a reference observable, enabling the systematic construction of recovery maps that preserve the latter. These recovery maps, although generally nonphysical, can be decomposed into realizable channels, enabling their application in noiseless expectation value estimation tasks. We provide explicit examples and compare our protocol with other error mitigation methods. We show that our protocol retrieves the noiseless expectation value of the reference observable and can achieve optimal sampling overhead, outperforming probabilistic error cancellation.

DOI: [10.1103/schg-nl3m](https://doi.org/10.1103/schg-nl3m)

I. INTRODUCTION

Due to unavoidable interactions with the environment, realistic quantum systems are inherently noisy, resulting in their evolution being governed by nonunitary quantum channels [1,2]. Consequently, any information we extract from the system is also corrupted by noise. A fundamental question that arises is then how to retrodict the initial state of a quantum system that has undergone such noisy dynamics. In particular, one might ask whether there exists a physical process (i.e., a quantum channel) that, starting from a state affected by noise, can recover the original state by undoing the effects of noise. However, in general, this is feasible only for unitary channels, as they are the only quantum channels that are reversible. Strictly speaking, the notion of time reversal in quantum theory is well defined for unitary evolutions only, that is, for closed systems' dynamics. A way to overcome this limitation is to introduce an additional element: a reference state, or prior, with respect to which the inversion is performed, along with the noise channel [3,4]. This leads to the introduction of a robust notion of time reversal for a generic quantum channel describing open quantum systems [5], in analogy with classical retrodiction.

In this framework, the reverse process is defined as a quantum map that only guarantees the recovery of the reference state. For instance, the well-known Petz recovery map [6–9] is a quantum channel that satisfies the aforementioned recovery property. If one further relaxes the requirement that the reverse process must be a quantum channel, other possible definitions of an inverse process for a given noise channel and reference

state become possible. Of course, if we are willing to relax the complete-positivity condition, the mathematical inverse of a channel can serve this purpose [10]. However, this inverse may not always be defined, nor be the most suitable choice, as we will discuss later.

The notion of a quantum state over time (QSOT) was introduced as the quantum analog of a joint probability distribution for a bipartite system across two timelike separated regions [11,12]. There are multiple ways to define states over time [3,11–16] because, while the postulates of quantum mechanics specify how to describe joint states in spacelike separated regions, they provide no guidance on how to jointly represent states that are timelike separated. In general, an appropriate definition of QSOT should be an operator acting on the tensor product of the Hilbert space at two different times, and if the two timelike separated states are causally related, then the QSOT should encode such correlations as well. Furthermore, as highlighted in Refs. [3,17], for a QSOT to capture temporal correlations, it must allow for negative eigenvalues, meaning that states over time are not positive operators in general. The authors of Ref. [5] then used the QSOT formalism and a notion of time-reversal symmetry in classical systems to systematically define Bayes' rule in the quantum framework. This approach results in an operational definition of “Bayesian inverse” of a quantum channel with respect to a reference state. Crucially, these maps inherently possess the recovery property of the reference state we have mentioned earlier. Naturally, different choices of QSOT lead to different recovery maps. For example, the Leifer-Spekkens QSOT [11,18] leads to a Bayesian inverse that corresponds to the Petz recovery map.

So far, most research efforts have been focused on defining retrodiction maps for a given noise channel and reference state. In this paper, we shift perspective and turn our attention to quantum observables. Our goal is to find recovery maps that protect specific observables from the effects of a given noisy dynamics. From a more operational standpoint, such

*Contact author: gabriele.bressanini@gmail.com

maps would allow us to devise protocols that preserve the expectation value of a specific observable against a given noise model, *for every possible state*. This is in contrast with state retrodiction, which cannot guarantee to perfectly infer the expectation values of observables beyond the reference state. This shift in perspective is motivated by the fact that several quantum information protocols and quantum algorithms rely on extracting information through expectation values, which are inevitably impacted by the noise affecting the system. Quantum error mitigation addresses the problem of protecting expectation values from noisy dynamics. In this context, unphysical recovery maps do not pose a problem, as they can be decomposed into physically realizable quantum channels [10,19,20], which further justifies the extension of the search for recovery maps to noncompletely positive (CP) maps. This quasiprobability decomposition (QPD) then serves as a simulation tool to estimate the noise-free expectation value, following the same strategy used in the well-known probabilistic error cancellation (PEC) [10,21] technique, as we will explore in more detail later. In Ref. [22], the authors address the problem of constructing recovery maps for specific observables under a given noise model using a different approach. The main difference between our methods lies in the fact that, while their approach requires solving a challenging optimization problem, we take a more fundamental route, leveraging time-reversal symmetries to systematically derive the recovery maps.

We summarize our main contributions in two perspectives as follows. First, from a foundational perspective, we generalize the state over time formalism to encompass quantum observables by introducing the concept of a quantum observable over time (QOOT). The latter is an operator that jointly describes two quantum observables that are timelike separated, while also capturing their temporal correlations. We show that the construction of the QOOT is not universally applicable, which makes a clear distinction from the QSOT, and we fully characterize the pairs of quantum channels and reference observables for which the QOOT can be properly defined. Second, we leverage our formulation of QOOT to implicitly and systematically define recovery maps via an equation that reflects a notion of time-reversal symmetry, thus offering a definition of time reversal for quantum channels with respect to a reference observable. On a more applied front, we investigate the potential of these recovery maps for quantum error mitigation, focusing on the task of estimating the noise-free expectation value of a specific reference observable subjected to a known noise channel. We explicitly construct two types of information recovery processes, namely, preprocessing and postprocessing protocols, both of which can be simulated on quantum devices using quasiprobability methods similar to PEC. We demonstrate the optimality of the sampling overhead required to simulate these protocols in certain cases. We note that, as being observable-specific recovery maps, the proposed protocols may become advantageous over existing approaches of implementing the inverse map of the noisy channel.

The rest of this paper is organized as follows. In Sec. II, we introduce the concept of quantum observable over time through its defining properties and present a necessary and sufficient condition under which such an object can be

properly defined. In Sec. III, we demonstrate how observables over time may be used to write an equation that expresses time-reversal symmetry of the system, leading to the implicit definition of recovery maps for both preprocessing and postprocessing protocols. In Sec. IV, we introduce a specific instance of the quantum observable over time, based on the Jordan product, which will be used throughout the paper due to its desirable properties. For this choice of QOOT, we then solve the equations, obtain the corresponding recovery maps, and study their properties. In Sec. V, we show how an unphysical recovery map, can be expressed in terms of physically implementable maps and discuss the application of this decomposition in noiseless expectation value estimation tasks. We then highlight differences and similarities between this approach and established error mitigation protocols. In Sec. VI, we explicitly compute recovery maps for relevant noise models, we decompose them in terms of physically realizable maps, and discuss their implementation cost, comparing it with known optimality results. Finally, in Sec. VII we draw conclusions, highlight remaining open questions, and offer final remarks.

II. QUANTUM OBSERVABLES OVER TIME

In this section, we introduce the quantum observable over time, i.e., the fundamental object that will be used throughout the paper to obtain recovery maps. Our construction follows that of quantum states over time presented in Ref. [5]. Let us consider an observable (i.e., a Hermitian operator) O and a Hermitian-preserving (HP) map $\mathcal{E}$ acting on the space of states as $\mathcal{E} : \rho \mapsto \mathcal{E}(\rho)$ (Schrödinger-like evolution). Equivalently, the adjoint map $\mathcal{E}^\dagger$ describes the dynamical evolution of observables (Heisenberg-like evolution), namely, $\mathcal{E}^\dagger : O \mapsto \mathcal{E}^\dagger(O)$. Here, O acts on the Hilbert space $\mathcal{H}_A$ and $\mathcal{E}^\dagger(O)$ acts on $\mathcal{H}_B$. Equivalently, we write $O \in \mathcal{L}(\mathcal{H}_A)$ and $\mathcal{E}^\dagger(O) \in \mathcal{L}(\mathcal{H}_B)$, where $\mathcal{L}(\mathcal{H})$ denotes the space of linear operators acting on the Hilbert space $\mathcal{H}$. The Schrödinger and Heisenberg pictures are well known to provide completely equivalent descriptions of a quantum system undergoing dynamical evolution. This duality becomes evident when considering the expectation value of an observable O for a system initially in state ρ , evolving under the map $\mathcal{E}$. Specifically, the expectation value can be expressed equivalently as

$$\text{Tr}[\mathcal{E}(\rho)O] = \text{Tr}[\rho \mathcal{E}^\dagger(O)]. \quad (1)$$

We seek to define a *quantum observable over time*, $\mathcal{E}^\dagger \star O$, i.e., an operator acting on $\mathcal{H}_A \otimes \mathcal{H}_B$ that can jointly describe the reference observable O and its dynamically evolved counterpart $\mathcal{E}^\dagger(O)$. In particular, such an object must correctly retrieve the marginal observables; namely, it should satisfy the following minimal requirements:

$$\text{Tr}_B[\mathcal{E}^\dagger \star O] = O, \quad (2)$$

$$\text{Tr}_A[\mathcal{E}^\dagger \star O] = \mathcal{E}^\dagger(O). \quad (3)$$

Additional desirable properties for $\mathcal{E}^\dagger \star O$ may be linearity in the reference observable, process linearity, and Hermiticity. We point out that the Leifer-Spekkens QSOT, which gives rise to the Petz recovery map, cannot be promoted to a QOOT. This is because its definition involves taking the square root of

TABLE I. Conditions on $\mathcal{E}$ and O under which the constructions of QOOT ($\mathcal{E}^\dagger \star O$), preprocessing ($\mathcal{P}$), and postprocessing ($\mathcal{R}$) protocols can be applied. The final condition for the postprocessing map ($\mathcal{R}$) should be verified manually after the map is obtained (see the main text for details).

	QOOT ($\mathcal{E}^\dagger \star O$)	Preprocessing ($\mathcal{P}$)	Postprocessing ($\mathcal{R}$)
Key property	$\text{Tr}_{\mathcal{B}}[\mathcal{E}^\dagger(O)] = O$ $\text{Tr}_{\mathcal{A}}[\mathcal{E}^\dagger(O)] = \mathcal{E}^\dagger(O)$	$O = \mathcal{P}^\dagger(\mathcal{E}^\dagger(O))$	$O = \mathcal{E}^\dagger(\mathcal{R}^\dagger(O))$
General condition(s)	$\text{Tr}[O] = \text{Tr}[\mathcal{E}^\dagger(O)]$	$\text{Tr}[O] = \text{Tr}[\mathcal{P}^\dagger(\mathcal{E}^\dagger(O))]$	$\mathcal{L}(\mathcal{H}_{\mathcal{B}}) = \mathcal{L}(\mathcal{H}_{\mathcal{A}})$ $\text{Tr}[O] = \text{Tr}[\mathcal{E}^\dagger(\mathcal{R}^\dagger(O))]$
Condition(s) for Jordan product	$\{O, \mathcal{E}(\mathcal{I})\} = 2O$	$\{\mathcal{E}^\dagger(O), \mathcal{P}(\mathcal{I})\} = 2\mathcal{E}^\dagger(O)$ (Automatically satisfied)	$\{O, \mathcal{R}(\mathcal{I})\} = 2O$ $\{\mathcal{R}^\dagger(O), \mathcal{E}(\mathcal{I})\} = 2\mathcal{R}^\dagger(O)$

the reference state. However, when the reference is a generic quantum observable, we cannot ensure its positivity. Interestingly, we show via a no-go theorem that, for a generic HP map $\mathcal{E}$ and quantum observable O , it is not always possible to construct a QOOT, i.e., an operator acting on $\mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}}$ satisfying the minimal partial trace conditions outlined above. This stands in sharp contrast to the QSOT formalism, where such a construction is always possible for any pair of reference state and quantum channel. We can easily see this by taking the partial trace over $\mathcal{H}_{\mathcal{B}}$ ($\mathcal{H}_{\mathcal{A}}$) of Eq. (2) [Eq. (3)], which leads to

$$\text{Tr}[\mathcal{E}^\dagger \star O] = \text{Tr}[O] = \text{Tr}[\mathcal{E}^\dagger(O)]. \quad (4)$$

Hence, a necessary and sufficient condition for the QOOT to be well defined is that the trace of $\mathcal{E}^\dagger(O)$ coincides with that of O (see also Table I). Notice that this does not automatically imply that the map $\mathcal{E}$ needs to be unital, since the above condition needs to only be satisfied for the reference observable O . In the following, we will assume that the combinations of maps and observables are such that QOOTs can be constructed. We note that, assuming Eq. (4) holds, the following trivial expression defines a valid QOOT:

$$\mathcal{E}^\dagger \star O = \frac{O \otimes \mathcal{E}^\dagger(O)}{\text{Tr}[O]}, \quad (5)$$

referred to as the *uncorrelated* quantum observable over time. However, this form has limited usefulness, as we will demonstrate in the next section, since it fails to encode any temporal correlations and treats O and $\mathcal{E}^\dagger(O)$ independently.

III. TIME-REVERSAL SYMMETRY AND QUANTUM OBSERVABLE RETRODICTION

In this section, we show how observables over time may be employed to implicitly define recovery maps via an equation that expresses time-reversal symmetry and resembles, in spirit, the Bayes rule and its quantum generalizations [5]. Let us consider a completely positive trace-preserving (CPTP) map $\mathcal{E} : \mathcal{L}(\mathcal{H}_{\mathcal{B}}) \rightarrow \mathcal{L}(\mathcal{H}_{\mathcal{A}})$ that models a noisy process, and a quantum observable $O \in \mathcal{L}(\mathcal{H}_{\mathcal{A}})$ that we wish to protect. We anticipate that, despite the limitations imposed by Eq. (4), the formalism developed in this paper is able to generate recovery maps for physically relevant scenarios, such as the protection of a Pauli observable under generalized amplitude damping (GAD) or stochastic Pauli noise.

For a preprocessing protocol, we seek an HP map $\mathcal{P} : \mathcal{L}(\mathcal{H}_{\mathcal{A}}) \rightarrow \mathcal{L}(\mathcal{H}_{\mathcal{B}})$ that satisfies the following (preprocessing)

recovery property:

$$O = \mathcal{P}^\dagger(\mathcal{E}^\dagger(O)). \quad (6)$$

The expression above leads to a *preprocessing* protocol because the action of $(\mathcal{P}^\dagger \circ \mathcal{E}^\dagger)$ on the observable O corresponds to the action of $(\mathcal{E} \circ \mathcal{P})$ on the state in Schrödinger picture. Hence, the preprocessing map $\mathcal{P}$ acts on the input state *before* it undergoes the noisy evolution. We can implicitly define such a map via the following equation, which reflects a certain time-reversal symmetry [5] we wish to impose (see Fig. 1 for a schematic representation):

$$\mathcal{E}^\dagger \star O = \tau(\mathcal{P}^\dagger \star \mathcal{E}^\dagger(O)). \quad (7)$$

Here, $\tau : \mathcal{L}(\mathcal{H}_{\mathcal{B}}) \otimes \mathcal{L}(\mathcal{H}_{\mathcal{A}}) \rightarrow \mathcal{L}(\mathcal{H}_{\mathcal{A}}) \otimes \mathcal{L}(\mathcal{H}_{\mathcal{B}})$ is the quantum-time-reversal map for observables over time, and it is defined as the unique conjugate-linear extension of

$$\tau(B \otimes A) = A^\dagger \otimes B^\dagger. \quad (8)$$

The operator swap introduced by τ places the two observables over time appearing in Eq. (7) on an equal footing: This is similar to what happens in the classical setting of the Bayes rule, where two joint probability distributions describing the same system are equated via time reversal [5]. Additionally, the map τ applies Hermitian conjugation, a feature that was shown to make the notion of a time-reversal map in the quantum setting more robust [5]. By taking the partial trace over Hilbert space $\mathcal{H}_{\mathcal{B}}$ on both sides of Eq. (7), we clearly obtain the preprocessing recovery property [Eq. (6)]. This implies that a solution to Eq. (7) always satisfies the recovery property for the system under study, provided that the two observables over time appearing in the equation are well defined, namely, when $\text{Tr}[O] = \text{Tr}[\mathcal{E}^\dagger(O)] = \text{Tr}[\mathcal{P}^\dagger(\mathcal{E}^\dagger(O))]$. We will come back to

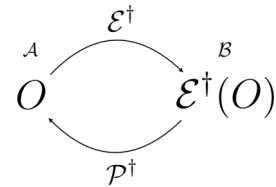


FIG. 1. Pictorial representation of the preprocessing protocol. Depending on which observable we take as reference to build the QOOT (time-reversal symmetry), the system may be equivalently described by $\mathcal{E}^\dagger \star O$ or by $\mathcal{P}^\dagger \star \mathcal{E}^\dagger(O)$. The preprocessing map acts on the noisy observable to retrieve the target observable O , i.e., $\mathcal{P}^\dagger(\mathcal{E}^\dagger(O)) = O$.

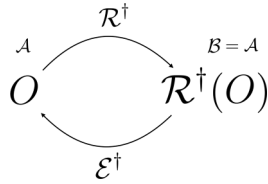


FIG. 2. Pictorial representation of the postprocessing protocol. Depending on which observable we take as reference to build the QOOT (time-reversal symmetry), the system may be equivalently described by $\mathcal{R}^\dagger \star O$ or by $\mathcal{E}^\dagger \star \mathcal{R}^\dagger(O)$. The recovery map acts on the target observable O so that the subsequent action of the noise map yields the target observable itself, i.e., $\mathcal{E}^\dagger(\mathcal{R}^\dagger(O)) = O$. We note that an additional condition $\mathcal{L}(\mathcal{H}_B) = \mathcal{L}(\mathcal{H}_A)$, simply referred to as $B = A$, is required for the postprocessing (see the main text for details).

this point in detail later, after choosing a specific functional form for the QOOT. The physical intuition behind Eq. (7) is that we may describe the same joint system using two observables over time: one for the “forward” process $\mathcal{E}^\dagger$, using O as the reference observable, and one for the “backward” process $\mathcal{P}^\dagger$ that instead utilizes the noisy observable $\mathcal{E}^\dagger(O)$ as reference operator. Note that if we use the uncorrelated QOOT from Eq. (5), Eq. (7) reduces to

$$O \otimes \mathcal{E}^\dagger(O) = \mathcal{P}^\dagger(\mathcal{E}^\dagger(O)) \otimes \mathcal{E}^\dagger(O), \quad (9)$$

which is solved by *any* map $\mathcal{P}$ such that $\mathcal{P}^\dagger(\mathcal{E}^\dagger(O)) = O$. However, this formulation does not indicate *which* of these recovery maps we should choose; hence, it does not provide a robust definition of recovery map.

Analogously, for a postprocessing protocol we seek a recovery map $\mathcal{R} : \mathcal{L}(\mathcal{H}_B) \rightarrow \mathcal{L}(\mathcal{H}_A)$ that satisfies the (postprocessing) recovery property

$$O = \mathcal{E}^\dagger(\mathcal{R}^\dagger(O)). \quad (10)$$

However, we note that $\mathcal{R}^\dagger : \mathcal{L}(\mathcal{H}_A) \rightarrow \mathcal{L}(\mathcal{H}_B)$ and $\mathcal{E}^\dagger : \mathcal{L}(\mathcal{H}_A) \rightarrow \mathcal{L}(\mathcal{H}_B)$, which implies $\mathcal{L}(\mathcal{H}_B) \subseteq \mathcal{L}(\mathcal{H}_A)$. Moreover, for information recovery, the image of the overall process $(\mathcal{E}^\dagger \circ \mathcal{R}^\dagger)$ should contain $\mathcal{L}(\mathcal{H}_A)$, which leads to $\mathcal{L}(\mathcal{H}_A) \subseteq \mathcal{L}(\mathcal{H}_B)$. Consequently, this requires an additional condition on the map $\mathcal{E}$ to have the same domain and image, i.e., $\mathcal{L}(\mathcal{H}_B) = \mathcal{L}(\mathcal{H}_A)$ to have a well-defined postprocessing protocol $\mathcal{R}$, which is a main difference from the preprocessing protocol $\mathcal{P}$. The expression above results in a *postprocessing* protocol, as the action of $(\mathcal{E}^\dagger \circ \mathcal{R}^\dagger)$ on the observable O corresponds to the action of $(\mathcal{R} \circ \mathcal{E})$ on the state in Schrödinger picture. Consequently, the recovery map $\mathcal{R}$ acts on the state *after* it has undergone the noisy dynamics. Such a recovery map may be implicitly defined via the following equation (see Fig. 2 for a schematic representation):

$$\mathcal{R}^\dagger \star O = \tau(\mathcal{E}^\dagger \star \mathcal{R}^\dagger(O)), \quad (11)$$

which, once again, expresses a time-reversal symmetry we impose on the joint system. Notice how the introduction of $\mathcal{R}^\dagger(O)$, i.e., the preprocessed target observable via the (adjoint) recovery map, is necessary in order to obtain a postprocessing protocol. Taking the partial trace of both sides

of Eq. (11) over $\mathcal{H}_B$ yields the recovery property [Eq. (10)], provided that $\text{Tr}[O] = \text{Tr}[\mathcal{R}^\dagger(O)] = \text{Tr}[\mathcal{E}^\dagger(\mathcal{R}^\dagger(O))]$.

IV. JORDAN PRODUCT QUANTUM OBSERVABLE OVER TIME

Previously, we introduced the quantum observable over time object in full generality and showed how it can be used to write Eqs. (7) and (11) that reflect a particular time-reversal symmetry of the system, and implicitly define recovery maps that protect the target observable O from noise modeled by the CPTP map $\mathcal{E}$. In this section, we present a specific instance of the quantum observable over time that will be used throughout the paper to derive recovery maps. To this end, we first introduce the Jamiołkowski state [23] of a quantum map $\mathcal{E}$, namely,

$$D[\mathcal{E}] = \sum_{ij} |i\rangle\langle j| \otimes \mathcal{E}(|j\rangle\langle i|), \quad (12)$$

where $\{|i\rangle\}$ denotes a generic basis of the Hilbert space. Given an HP map $\mathcal{E}^\dagger : \mathcal{L}(\mathcal{H}_A) \rightarrow \mathcal{L}(\mathcal{H}_B)$ and a reference observable O acting on $\mathcal{H}_A$, we define the Jordan product quantum observable over time as

$$\mathcal{E}^\dagger \star O = \frac{1}{2}\{O \otimes \mathcal{I}, D[\mathcal{E}^\dagger]\}, \quad (13)$$

where $\mathcal{I}$ denotes the identity operator and the curly brackets denote the usual anticommutator. Substituting the definition of the Jamiołkowski state into the equation above leads to

$$\mathcal{E}^\dagger \star O = \frac{1}{2} \sum_{ij} \{O, |i\rangle\langle j|\} \otimes \mathcal{E}^\dagger(|j\rangle\langle i|). \quad (14)$$

This particular form of the QOOT, inspired by the Jordan product state over time in Ref. [5], has desirable properties, including linearity in the reference observable, linearity in the quantum map, and it is a Hermitian operator. We can now revisit the necessary and sufficient conditions for the Jordan product QOOT to be well defined; i.e., we have to make sure that the defining properties Eqs. (2) and (3) are satisfied. In particular, taking the partial trace of Eq. (13) over $\mathcal{H}_B$ leads to

$$\begin{aligned} \text{Tr}_B[\mathcal{E}^\dagger \star O] &= \frac{1}{2} \sum_{ij} \{O, |i\rangle\langle j|\} \text{Tr}[\mathcal{E}^\dagger(|j\rangle\langle i|)] \\ &= \frac{1}{2} \sum_{ij} \{O, |i\rangle\langle j|\} \langle i| \mathcal{E}(\mathcal{I}) |j\rangle \\ &= \frac{1}{2} \{O, \mathcal{E}(\mathcal{I})\}, \end{aligned} \quad (15)$$

where we have used the definition of adjoint map, namely, $\text{Tr}[\mathcal{E}^\dagger |j\rangle\langle i|] = \text{Tr}[|j\rangle\langle i| \mathcal{E}(\mathcal{I})]$. For Eq. (2) to be satisfied, we need

$$\{O, \mathcal{E}(\mathcal{I})\} = 2O \quad (16)$$

to hold (see also Table I). Clearly, if $\mathcal{E}$ is unital, the condition is satisfied; however, that is not the only possibility. In fact, one easily shows that the necessary and sufficient condition for Eq. (16) to hold is

$$\mathcal{E}(\mathcal{I}) = \mathcal{I} + \bar{O}, \quad (17)$$

where $\bar{O}$ is a generic operator that anticommutes with O . Analogously, the partial trace of the Jordan product QOOT over $\mathcal{H}_A$ gives

$$\begin{aligned}\text{Tr}_A[\mathcal{E}^\dagger \star O] &= \frac{1}{2} \sum_{ij} \text{Tr}[\{O, |i\rangle\langle j|\}] \mathcal{E}^\dagger(|j\rangle\langle i|) \\ &= \sum_{ij} \langle j| O |i\rangle \mathcal{E}^\dagger(|j\rangle\langle i|) \\ &= \mathcal{E}^\dagger(O),\end{aligned}\quad (18)$$

where in the last equality we used the linearity of $\mathcal{E}^\dagger$. Notice that Eq. (3) is in this case automatically satisfied. We are now ready to substitute the expression of the Jordan product QOOT into Eqs. (7) and (11) and solve them for $\mathcal{P}$ and $\mathcal{R}$, respectively.

A. Preprocessing protocol

Let us consider a reference observable O acting on $\mathcal{H}_A$ that we want to protect from a noisy process modeled by a CPTP map $\mathcal{E} : \mathcal{L}(\mathcal{H}_B) \rightarrow \mathcal{L}(\mathcal{H}_A)$. As outlined in Sec. III, a preprocessing map $\mathcal{P} : \mathcal{L}(\mathcal{H}_A) \rightarrow \mathcal{L}(\mathcal{H}_B)$ is obtained by solving

$$\mathcal{E}^\dagger \star O = \tau(\mathcal{P}^\dagger \star \mathcal{E}^\dagger(O)). \quad (19)$$

Substituting the expression of the Jordan product QOOT into the previous equation yields

$$\tau(\{O \otimes \mathcal{I}, D[\mathcal{E}^\dagger]\}) = \{\mathcal{I} \otimes \mathcal{E}^\dagger(O), D[\mathcal{P}^\dagger]\}, \quad (20)$$

where we have used the fact that applying τ twice corresponds to the identity channel. The action of the time-reversal operator τ on the left-hand side of Eq. (20) reads

$$\begin{aligned}\tau(\{O \otimes \mathcal{I}, D[\mathcal{E}^\dagger]\}) &= \{\tau(O \otimes \mathcal{I}), \tau(D[\{\mathcal{E}^\dagger\}])\} \\ &= \{\mathcal{I} \otimes O, D[\mathcal{E}]\},\end{aligned}\quad (21)$$

where we have used the fact that O is Hermitian and that $\tau(D[\mathcal{E}^\dagger]) = D[\mathcal{E}]$ (see Appendix A for a detailed derivation of this identity). Hence, we obtain the following equation:

$$\{\mathcal{I} \otimes O, D[\mathcal{E}]\} = \{\mathcal{I} \otimes \mathcal{E}^\dagger(O), D[\mathcal{P}^\dagger]\}. \quad (22)$$

As already mentioned in the previous section, for the two Jordan product QOOTs appearing in the equation above to be well defined, we need the following two conditions to be satisfied (see also Table I):

$$\{O, \mathcal{E}(\mathcal{I})\} = 2O \iff \mathcal{E}(\mathcal{I}) = \mathcal{I} + \bar{O}, \quad (23)$$

$$\{\mathcal{E}^\dagger(O), \mathcal{P}(\mathcal{I})\} = 2\mathcal{E}^\dagger(O) \iff \mathcal{P}(\mathcal{I}) = \mathcal{I} + \overline{\mathcal{E}^\dagger(O)}. \quad (24)$$

The condition given in Eq. (23) may be directly checked using the input of the problem, namely, O and $\mathcal{E}$. On the other hand, in what follows, we show that Eq. (24) is automatically satisfied by any solution $\mathcal{P}$ to Eq. (22). Let us now consider the eigendecomposition of the noisy reference observable, namely,

$$\mathcal{E}^\dagger(O) = \sum_k q_k |\omega_k\rangle\langle\omega_k|, \quad (25)$$

where for now we assume that the eigenvalues $\{q_k\}$ are such that $q_k + q_j \neq 0$ for all j and k (in Sec. VI, we show how

an appropriate regularization of the reference observable may be employed to address the cases where this condition is not satisfied). In Appendix B, we show that the solution to Eq. (22) reads

$$\mathcal{P}^\dagger(|\omega_k\rangle\langle\omega_\ell|) = (q_k + q_\ell)^{-1} \{O, \mathcal{E}(|\omega_k\rangle\langle\omega_\ell|)\}. \quad (26)$$

This map is clearly HP, and we can easily show that $\mathcal{P}^\dagger$ is TP as well (i.e., $\mathcal{P}$ is unital), in fact

$$\begin{aligned}\text{Tr}[\mathcal{P}^\dagger(|\omega_k\rangle\langle\omega_\ell|)] &= (q_k + q_\ell)^{-1} \text{Tr}[\{O, \mathcal{E}(|\omega_k\rangle\langle\omega_\ell|)\}] \\ &= 2(q_k + q_\ell)^{-1} \langle\omega_\ell| \mathcal{E}^\dagger(O) |\omega_k\rangle \\ &= 2(q_k + q_\ell)^{-1} q_k \delta_{k\ell} = \delta_{k\ell},\end{aligned}\quad (27)$$

where $\delta_{k\ell}$ denotes the Kronecker delta. A consequence of the fact that $\mathcal{P}^\dagger$ is HPTP is that can be expressed as a linear combination of CPTP maps. Additionally, in Appendix C we prove the desirable property that if the map $\mathcal{E}$ corresponds to a unitary evolution (i.e., it is invertible and its inverse is CPTP), then $\mathcal{E}^{-1}$ is indeed the preprocessing map $\mathcal{P}$.

B. Postprocessing protocol

As previously outlined in Sec. III, a recovery map $\mathcal{R} : \mathcal{L}(\mathcal{H}_B) \rightarrow \mathcal{L}(\mathcal{H}_A)$ for the postprocessing protocol [we recall that $\mathcal{L}(\mathcal{H}_B) = \mathcal{L}(\mathcal{H}_A)$ is required for this case] is obtained by solving

$$\mathcal{R}^\dagger \star O = \tau(\mathcal{E}^\dagger \star \mathcal{R}^\dagger(O)). \quad (28)$$

Substituting the expression of the Jordan product QOOT into the equation above yields

$$\{O \otimes \mathcal{I}, D[\mathcal{R}^\dagger]\} = \tau(\{\mathcal{I} \otimes \mathcal{R}^\dagger(O), D[\mathcal{E}^\dagger]\}). \quad (29)$$

Analogously to the preprocessing protocol, for the two Jordan product QOOTs appearing in Eq. (29) to be well defined, we need the following conditions to hold (see also Table I):

$$\{O, \mathcal{R}(\mathcal{I})\} = 2O \iff \mathcal{R}(\mathcal{I}) = \mathcal{I} + \bar{O}, \quad (30)$$

$$\{\mathcal{R}^\dagger(O), \mathcal{E}(\mathcal{I})\} = 2\mathcal{R}^\dagger(O) \iff \mathcal{E}(\mathcal{I}) = \mathcal{I} + \overline{\mathcal{R}^\dagger(O)}. \quad (31)$$

Coming back to Eq. (29), the action of the time-reversal operator τ reads

$$\begin{aligned}\tau(\{\mathcal{I} \otimes \mathcal{R}^\dagger(O), D[\mathcal{E}^\dagger]\}) &= \{\tau(\mathcal{I} \otimes \mathcal{R}^\dagger(O)), \tau(D[\mathcal{E}^\dagger])\} \\ &= \{\mathcal{R}^\dagger(O) \otimes \mathcal{I}, D[\mathcal{E}]\},\end{aligned}\quad (32)$$

where we have used the fact that we look for solutions $\mathcal{R}$ to the equation that are HP, O is Hermitian, and $\tau(D[\mathcal{E}^\dagger]) = D[\mathcal{E}]$. Hence, we obtain

$$\{O \otimes \mathcal{I}, D[\mathcal{R}^\dagger]\} = \{\mathcal{R}^\dagger(O) \otimes \mathcal{I}, D[\mathcal{E}]\}. \quad (33)$$

The difficulty in solving this equation lies in the fact that, unlike the preprocessing protocol, the recovery map now appears on both sides of the equation and, more importantly, we do not know anything about $\mathcal{R}^\dagger(O)$ *a priori*. In the following, we briefly outline the techniques used to solve Eq. (33), while a detailed derivation can be found in Appendix D. First, let us consider the eigendecomposition of the reference observable, i.e., $O = \sum_i q_i |\omega_i\rangle\langle\omega_i|$. Notice that, while we do not know what $\mathcal{R}^\dagger(O)$ is, we may express it as $\mathcal{R}^\dagger(O) = \sum_i q_i \mathcal{R}^\dagger(|\omega_i\rangle\langle\omega_i|)$, by linearity. Projecting both

sides of Eq. (33) onto the joint Hilbert space basis $\{|\omega_i\rangle \otimes |\omega_j\rangle\}$ yields

$$(q_k + q_\ell)\mathcal{R}^\dagger(|\omega_\ell\rangle\langle\omega_k|) = \sum_i q_i \{\mathcal{R}^\dagger(|\omega_i\rangle\langle\omega_i|), \mathcal{E}(|\omega_\ell\rangle\langle\omega_k|)\}. \quad (34)$$

The equation above can be recast as

$$X_{k\ell} = \sum_i \{X_{ii}, A_{k\ell}\}, \quad (35)$$

where $X_{k\ell} = (q_k + q_\ell)\mathcal{R}^\dagger(|\omega_k\rangle\langle\omega_\ell|)$, and $A_{k\ell} = \mathcal{E}(|\omega_k\rangle\langle\omega_\ell|)/2$. The HP requirement for the recovery map is equivalent to enforcing $X_{k\ell}^\dagger = X_{\ell k}$. Finally, we need to make sure that the conditions necessary and sufficient for the QOOTs to be well defined are satisfied. In particular, Eq. (31) can be recast as $0 = \sum_i \{X_{ii}, \mathcal{E}(\mathcal{I}) - \mathcal{I}\}$. A unital noise map always satisfies this constraint; however, in general, this would impose an additional constraint on the recovery map $\mathcal{R}$. Similarly, one can show (see Appendix F) that Eq. (30) is equivalent to imposing the additional constraint $\text{Tr}[X_{k\ell}] = q_k \delta_{k\ell}$. Notice how this is different from the preprocessing scenario, where the arising recovery maps $\mathcal{P}$ are automatically unital and hence satisfy Eq. (24) without needing to impose additional constraints.

Let us now consider the case of a qubit system, meaning that we can decompose all operators appearing in the equation in the Pauli basis, namely,

$$X_{k\ell} = \sum_{j=0}^3 x_j^{(k\ell)} \sigma_j, \quad A_{k\ell} = \sum_{j=0}^3 a_j^{(k\ell)} \sigma_j, \quad (36)$$

where $\{\sigma_0, \sigma_1, \sigma_2, \sigma_3\} \equiv \{\mathcal{I}, X, Y, Z\}$ denote the Pauli operators. We can now substitute these expressions into Eq. (35) and obtain an equivalent system of linear equations for the eight coefficients $x_j^{(k\ell)}$:

$$x_0^{(kl)} = 2 \sum_{i,j} x_j^{(ii)} a_j^{(kl)}, \quad (37)$$

$$x_j^{(kl)} = 2 \sum_i (x_0^{(ii)} a_j^{(kl)} + x_j^{(ii)} a_0^{(kl)}), \quad j \in \{1, 2, 3\}. \quad (38)$$

In Appendix E, we outline a detailed derivation of the system of equations above and a generalization to d -dimensional systems. Analogously to the preprocessing case, if $\mathcal{E}$ is a unitary evolution (i.e., the inverse exists and it is CPTP), then $\mathcal{E}^{-1}$ is a recovery map $\mathcal{R}$ that solves Eq. (33), as can be shown by direct substitution.

V. SIMULATION OF RECOVERY MAPS AND APPLICATION TO ERROR MITIGATION

Physically implementable operations are described using CPTP maps. More generally, if we allow these operations to have a probabilistic rather than deterministic implementation, they can be characterized by completely positive, trace nonincreasing (CPTN) maps. Nevertheless, nonphysical maps find application in several quantum information processing tasks. In fact, while an HP—but non-CP—map $\mathcal{R}$ represents a physically unattainable quantum operation, it is always possible to

express it as a linear combination of CPTN maps $\{\mathcal{F}_i\}$, namely,

$$\mathcal{R} = \sum_i c_i \mathcal{F}_i, \quad (39)$$

where c_i are real (possibly negative) coefficients. In particular, we might think of the total negativity as a measure of the “unphysicality” of the map. We will refer to Eq. (39) as a QPD of $\mathcal{R}$. If $\mathcal{R}$ is also TP, it turns out that we can always decompose it in terms of CPTP maps and that $\sum_i c_i = 1$. A QPD of an unphysical map $\mathcal{R}$ may be used as a simulation tool to estimate expectation values of the form $\text{Tr}[\mathcal{R}(\rho)O]$. In fact, we may express the latter as

$$\begin{aligned} \text{Tr}[\mathcal{R}(\rho)O] &= \sum_i c_i \text{Tr}[\mathcal{F}_i(\rho)O] \\ &= \gamma \sum_i \text{sgn}(c_i) p_i \text{Tr}[\mathcal{F}_i(\rho)O], \end{aligned} \quad (40)$$

where $\gamma = \sum_i |c_i|$ and $p_i = |c_i|/\gamma$ is a probability distribution. The idea is now to sample a quantum channel $\mathcal{F}^{(j)}$ from $\{\mathcal{F}_i\}$ according to the probability distribution p_i , apply it to the state ρ , and subsequently measure the observable O and record the measurement outcome $o^{(j)}$ (j labels the sample number). After N sampling rounds, we can construct the following unbiased estimator for $\text{Tr}[\mathcal{R}(\rho)O]$:

$$E = \frac{\gamma}{N} \sum_{j=1}^N \text{sgn}(c^{(j)}) o^{(j)}. \quad (41)$$

Furthermore, from Hoeffding’s inequality [24] the number of samples required for E to be ϵ close to the true value $\text{Tr}[\mathcal{R}(\rho)O]$, with probability no less than $1 - \delta$, scales as $N \propto \gamma^2 \log(2/\delta)/\epsilon^2$ [25]. The parameter γ can then be used as a measure of the *simulation cost* of $\mathcal{R}$, according to the decomposition Eq. (39) (see also Fig. 3 for an illustration of the overall protocol). It is clear that different decompositions of the same unphysical map $\mathcal{R}$ lead to different γ and different simulation costs as a result. In Ref. [20], the authors show that the optimal decomposition (i.e., achieving the minimum value of γ) can be computed via semidefinite programming.

A prominent example of the simulation technique described above is provided by probabilistic error cancellation [10], an error mitigation technique aimed at producing bias-free expectation values for systems that are subject to a known noise model, characterized by the CPTP map $\mathcal{E}$. Assuming that the noise map can be mathematically inverted, the key insight behind PEC is that we can simulate $\mathcal{E}^{-1} = \sum_i c_i \mathcal{F}_i$ (in the sense discussed above, since the inverse of a CPTP map is HPTP) and obtain bias-free estimates of the noiseless expectation value $\text{Tr}[\rho O]$ by performing measurements of O on $\mathcal{F}_i(\mathcal{E}(\rho))$ and by classical postprocessing of data. Within this framework, the simulation cost we mentioned earlier can also be interpreted as *sampling overhead*, as it can be shown [10,26] that the cost associated with obtaining a bias-free estimate of the noiseless expectation value is an increase in the variance of the estimator. In particular, the ratio between the variance of the bias-free estimator and that of the noisy estimator can be shown to scale with γ^2 .

Inspired by PEC, our recovery maps may also be employed to obtain noise-free expectation values of a quantum

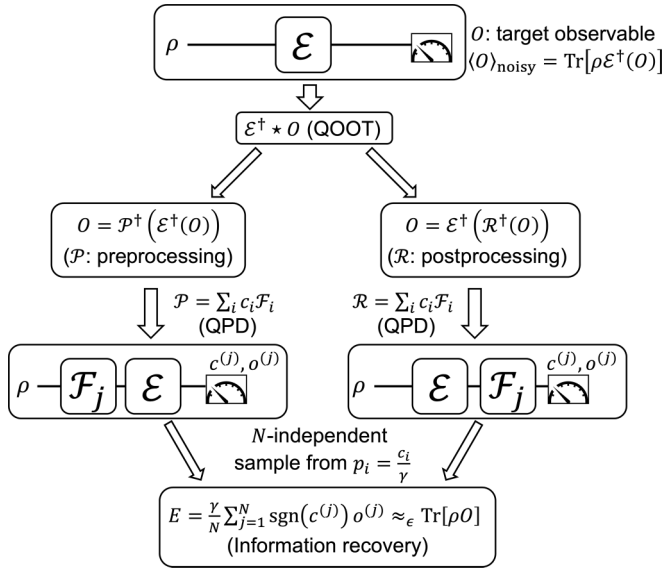


FIG. 3. Schematic illustrating the simulation of recovery maps for constructing an unbiased estimator E of the noiseless observable $\text{Tr}[\rho O]$ from measurements of the noisy observable $\text{Tr}[\rho \mathcal{E}^\dagger(O)]$. The number of samples N to achieve ϵ error scales as γ^2 , where $\gamma = \sum_i |c_i|$ (see the main text for details).

observable O subject to a CPTP noise map $\mathcal{E}$. In fact, by construction, we have that

$$\text{Tr}[\rho O] = \text{Tr}[\rho \mathcal{E}^\dagger(\mathcal{R}^\dagger(O))], \quad (42)$$

for the postprocessing protocol, and

$$\text{Tr}[\rho O] = \text{Tr}[\rho \mathcal{P}^\dagger(\mathcal{E}^\dagger(O))], \quad (43)$$

for the preprocessing protocol. All we need to do is compute the appropriate map ($\mathcal{R}$ or $\mathcal{P}$) as described in Sec. IV, and decompose it in terms of physically implementable maps. Our protocols might prove particularly useful when the inverse of the noise channel does not exist; hence, the framework of PEC is not applicable. Furthermore, our protocols would also be useful in situations where the implementation cost of our recovery maps outperforms that of the inverse noise map.

VI. EXAMPLES

In this section, we provide explicit constructions of recovery maps (preprocessing and postprocessing protocols) for Pauli reference observables and physically relevant noise models, namely, GAD and stochastic Pauli noise. The latter is of particular importance, since Pauli twirling [27] may be employed to convert an arbitrary noise channel into a stochastic Pauli noise channel. For all examples, we provide a decomposition of the recovery maps in terms of physically implementable channels and compute the cost of implementation γ . Comparing the latter with known optimality results reported in Ref. [22] reveals that the recovery maps presented here and their decompositions are optimal.

A. GAD channel and Pauli X (preprocessing)

The generalized amplitude damping channel can be viewed as the qubit analog of the bosonic thermal channel [28], and

may be expressed as $\mathcal{E}(\rho) = \sum_{j=1}^4 K_j \rho K_j^\dagger$, where the Kraus operators read

$$K_1 = \sqrt{p}|0\rangle\langle 0| + \sqrt{p(1-\epsilon)}|1\rangle\langle 1|, \quad (44)$$

$$K_2 = \sqrt{p\epsilon}|0\rangle\langle 1|, \quad (45)$$

$$K_3 = \sqrt{(1-p)(1-\epsilon)}|0\rangle\langle 0| + \sqrt{(1-p)}|1\rangle\langle 1|, \quad (46)$$

$$K_4 = \sqrt{\epsilon(1-p)}|1\rangle\langle 0|, \quad (47)$$

with $0 \leq p \leq 1$ and $0 \leq \epsilon \leq 1$. One easily shows that the GAD channel is nontrivially unital only for $p = 1/2$ and that the amplitude damping channel can be retrieved by setting $p = 1$. In this example, we set the reference observable O to be the Pauli X operator, and look for a preprocessing map $\mathcal{P}$, i.e., satisfying $\mathcal{P}^\dagger(\mathcal{E}^\dagger(X)) = X$. Notice that the necessary and sufficient condition in Eq. (23) for the quantum observable over time (hence, the task) to be well defined is satisfied, since $\mathcal{E}(\mathcal{I}) = \mathcal{I} + \epsilon(2p-1)Z$ and $\{X, Z\} = 0$. In Sec. IV A, we showed that the expression for such a recovery map reads

$$\mathcal{P}^\dagger(|\omega_k\rangle\langle\omega_\ell|) = (q_k + q_\ell)^{-1} \{O, \mathcal{E}(|\omega_k\rangle\langle\omega_\ell|)\}, \quad (48)$$

where q_k and $|\omega_k\rangle$ are the eigenvalues and eigenstates of the noisy reference observable $\mathcal{E}^\dagger(O)$, respectively. We remind the reader that the expression above is formally valid for $q_k + q_\ell \neq 0$. However, this condition is not satisfied in our scenario, since $\mathcal{E}^\dagger(X) = \sqrt{1-\epsilon}X$ and X is traceless. We may work around this potential setback by *regularizing* the reference observable; i.e., we define

$$\tilde{X} = \begin{pmatrix} \lambda & 1 \\ 1 & \lambda \end{pmatrix} = \lambda \mathcal{I} + X, \quad (49)$$

where λ is a positive regularization parameter that will be taken to the limit of zero at the end of the calculations. The action of the (adjoint) GAD channel on $\tilde{X}$ reads

$$\mathcal{E}^\dagger(\tilde{X}) = \lambda \mathcal{I} + \sqrt{1-\epsilon}X = q_1|\omega_1\rangle\langle\omega_1| + q_2|\omega_2\rangle\langle\omega_2|, \quad (50)$$

where

$$q_1 = \lambda - \sqrt{1-\epsilon}, \quad |\omega_1\rangle = \frac{-|0\rangle + |1\rangle}{\sqrt{2}} \equiv -|-\rangle, \quad (51)$$

$$q_2 = \lambda + \sqrt{1-\epsilon}, \quad |\omega_2\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \equiv |+\rangle. \quad (52)$$

We can now compute the preprocessing map via Eq. (48) and, after some algebra, obtain

$$\mathcal{P}^\dagger(\mathcal{I}) = \mathcal{I}, \quad (53)$$

$$\mathcal{P}^\dagger(X) = \frac{X}{\sqrt{1-\epsilon}}, \quad (54)$$

$$\mathcal{P}^\dagger(Y) = \sqrt{1-\epsilon}Y, \quad (55)$$

$$\mathcal{P}^\dagger(Z) = (1-\epsilon)Z, \quad (56)$$

where we have already taken the limit $\lambda \rightarrow 0$ [the potentially problematic $1/\lambda$ dependency of the terms $\mathcal{P}^\dagger(|+\rangle\langle-|)$ and $\mathcal{P}^\dagger(|-\rangle\langle+|)$ actually cancels out with a λ factor coming out of the numerator]. One can easily check that this map indeed satisfies the recovery property, i.e., $\mathcal{P}^\dagger(\mathcal{E}^\dagger(X)) = X$. Interestingly, we note that although $\mathcal{E}^\dagger$ and $\mathcal{P}^\dagger$ do not commute in

general, it is true that

$$X = \mathcal{P}^\dagger(\mathcal{E}^\dagger(X)) = \mathcal{E}^\dagger(\mathcal{P}^\dagger(X)). \quad (57)$$

Hence, it follows that the recovery map above may also be employed in a postprocessing protocol.

The next step is to find an operator-sum (i.e., Kraus-like) representation of the recovery map we just found. In particular, we want to express $\mathcal{P}$ as a linear combination of physically realizable CP channels. A common technique to achieve a Kraus-like decomposition of $\mathcal{P}$, is diagonalizing the recovery map's Choi matrix, namely,

$$\mathcal{J}[\mathcal{P}] = \sum_{ij} \mathcal{P}(|i\rangle\langle j|) \otimes |i\rangle\langle j|, \quad (58)$$

followed by the unvectorization of its eigenvectors. After some algebra, we obtain the following operator-sum decomposition:

$$\mathcal{P}(\sigma) = \sum_{i=1}^4 e_i E_i \sigma E_i^\dagger, \quad (59)$$

where $\{e_i\}$ are the Choi matrix eigenvalues, namely,

$$e_1 = (2 - \epsilon)\alpha_+, \quad e_2 = (2 - \epsilon)\alpha_-, \quad e_3 = \epsilon\alpha_+, \quad e_4 = \epsilon\alpha_-, \quad (60)$$

with $\alpha_\pm = \frac{1}{2} \pm \frac{1}{2\sqrt{1-\epsilon}}$. Notice that e_2 and e_4 are negative, meaning that the Choi matrix of the recovery map is not positive semidefinite; hence, the recovery map itself is not CP. The operators E_i are obtained from the unvectorization of the Choi matrix (normalized) eigenvectors, in particular, we find

$$E_1 = \frac{\mathcal{I}}{\sqrt{2}}, \quad E_2 = \frac{Z}{\sqrt{2}}, \quad E_3 = -\frac{X}{\sqrt{2}}, \quad E_4 = \frac{iY}{\sqrt{2}}. \quad (61)$$

Hence, the recovery map can be decomposed in terms of CPTP maps and, in particular, takes the form of a stochastic Pauli map with some negative coefficients. As previously explained, the cost of implementation of an HPTP map, given a specific linear decomposition in terms of CPTP maps, may be evaluated by summing the absolute value of said coefficients. For the above operator-sum representation of $\mathcal{P}$, the cost of implementation is equal to

$$\gamma_{\mathcal{P}} = \frac{1}{2}(|e_1| + |e_2| + |e_3| + |e_4|) = \frac{1}{\sqrt{1-\epsilon}}. \quad (62)$$

We note that the decomposition in Eq. (59), obtained from the eigendecomposition of the Choi matrix, need not necessarily be optimal in general. However, for a mixed unitary HPTP map $\mathcal{T}$, represented as

$$\mathcal{T}(\cdot) = \sum_i r_i U_i(\cdot) U_i^\dagger, \quad (63)$$

with $r_i \in \mathbb{R}$ and $\sum_i r_i = 1$, such a decomposition with cost $\sum_i |r_i|$ is proven to be optimal in Theorem 13 of Ref. [20]. As the GAD channel falls into this case by noting that E_i in Eq. (59) are Pauli matrices, the cost given in Eq. (62) becomes optimal.

We may first compare this with the *optimal* (i.e., minimized over all possible decompositions in terms of linear combinations of CPTP maps) implementation cost of the GAD channel

inverse map. In Ref. [20] the authors show that the latter reads

$$\gamma_{\mathcal{E}^{-1}} = \frac{|1 - 2p|\epsilon + 1}{1 - \epsilon}. \quad (64)$$

One then shows that $\gamma_{\mathcal{P}} < \gamma_{\mathcal{E}^{-1}}$ for all possible values of p and ϵ , meaning that the sampling overhead associated with our recovery map with the decomposition above outperforms the traditional PEC protocol relying on implementing the channel inverse.

For a given noise channel and reference observable, the authors of Ref. [22] employed semidefinite programming to optimize the implementation cost across all possible Hermitian-preserving trace-scaling recovery maps and across their decompositions in terms of CPTP maps. In particular, for the GAD channel and Pauli X reference observable, they demonstrated that the implementation cost is lower bounded by Eq. (62), thus making the recovery map we found *optimal* for this task.

B. Stochastic Pauli channel and Pauli Z (postprocessing)

Another key example is the stochastic Pauli channel, defined as

$$\mathcal{E}(\rho) = p_0 \rho + p_1 X \rho X + p_2 Y \rho Y + p_3 Z \rho Z, \quad (65)$$

where the coefficients p_i constitute a probability distribution, satisfying $p_i \geq 0$ and $\sum_i p_i = 1$. Equation (65) encompasses several widely employed noise models. For instance, setting $p_1 = p_2 = p_3 = \frac{\lambda}{4}$ and $p_0 = 1 - \frac{3\lambda}{4}$ yields the depolarizing channel, while choosing $p_0 = 1 - p$, $p_1 = p_2 = 0$, $p_3 = p$ for $0 < p < 1$ results in the dephasing channel. Furthermore, the stochastic Pauli channel is of particular interest because an arbitrary single-qubit noise channel may be reduced to a stochastic Pauli noise channel through Pauli twirling [27]. In this example, we set the reference observable O to be the Pauli Z operator (analogous results would follow for X or Y) and seek a postprocessing recovery map $\mathcal{R}$, hence satisfying $\mathcal{E}^\dagger(\mathcal{R}^\dagger(Z)) = Z$. Since any stochastic Pauli channel is unital, the consistency condition in Eq. (31) is automatically satisfied. On the other hand, the remaining condition in Eq. (30) still needs to be imposed. We have previously demonstrated that a postprocessing recovery map can be obtained by solving the following equation:

$$(q_k + q_\ell) \mathcal{R}^\dagger(|\omega_\ell\rangle\langle\omega_k|) = \sum_i q_i \{\mathcal{R}^\dagger(|\omega_i\rangle\langle\omega_i|), \mathcal{E}(|\omega_\ell\rangle\langle\omega_k|)\}, \quad (66)$$

where q_k and $|\omega_k\rangle$ are the eigenvalues and corresponding eigenstates of the reference observable, respectively. Similar to the previous example, we regularize the reference observable by defining $\tilde{Z} = Z + \lambda \mathcal{I}$, where $\lambda \geq 0$ serves as a real regularization parameter that will be taken to zero at the end of the calculation. The eigendecomposition of $\tilde{Z}$ is thus characterized by

$$q_1 = 1 + \lambda, \quad |\omega_1\rangle = |0\rangle, \quad (67)$$

$$q_2 = -1 + \lambda, \quad |\omega_2\rangle = |1\rangle. \quad (68)$$

As shown in Sec. IV, computing the recovery map amounts to solving the system of linear equations given by Eqs. (37)

and (38) with the coefficients $a_j^{(k\ell)}$ reading

$$a_0^{(k\ell)} = \frac{1}{2}\omega_0^{(k\ell)}, \quad a_j^{(k\ell)} = (p_0 + p_j - \frac{1}{2})\omega_j^{(k\ell)}. \quad (69)$$

Here, $\omega_j^{(k\ell)}$ are the Pauli basis coefficients for the outer product of the basis elements, namely,

$$|\omega_k\rangle\langle\omega_\ell| = \sum_j \omega_j^{(k\ell)} \sigma_j. \quad (70)$$

The solution to these equations, upon taking $\lambda \rightarrow 0$, yields the following recovery map:

$$\begin{aligned} \mathcal{R}^\dagger(\mathcal{I}) &= \mathcal{I}, \\ \mathcal{R}^\dagger(X) &= (p_0 + p_1 - p_2 - p_3)X, \\ \mathcal{R}^\dagger(Y) &= (p_0 - p_1 + p_2 - p_3)Y, \\ \mathcal{R}^\dagger(Z) &= \frac{1}{p_0 - p_1 - p_2 + p_3}Z. \end{aligned} \quad (71)$$

The recovery map $\mathcal{R}$ is thus well defined for all stochastic Pauli channels except for the particular cases of $p_0 + p_3 = \frac{1}{2}$. Furthermore, since $\mathcal{R}^\dagger$ is manifestly TP, it implies that $\mathcal{R}$ is unital, thereby satisfying the consistency condition in Eq. (30). Next, we want to express the recovery map as a linear combination of physically realizable maps. To this end, following the procedure outlined in the previous example, we compute the Choi matrix of $\mathcal{R}$, diagonalize it, and unvectorize its (normalized) eigenvectors. After some algebra, we obtain the following operator-sum representation:

$$\begin{aligned} \mathcal{R}_Z(\sigma) &= \frac{\alpha_0 + \beta_0}{2}\sigma + \frac{\alpha_1 + \beta_1}{2}X\sigma X \\ &\quad + \frac{\alpha_1 - \beta_1}{2}Y\sigma Y + \frac{\alpha_0 - \beta_0}{2}Z\sigma Z, \end{aligned} \quad (72)$$

where

$$\begin{aligned} \alpha_0 &= \frac{p_0 + p_3}{p_0 - p_1 - p_2 + p_3}, \quad \beta_0 = p_0 - p_3, \\ \alpha_1 &= -\frac{p_1 + p_2}{p_0 - p_1 - p_2 + p_3}, \quad \beta_1 = p_1 - p_2. \end{aligned} \quad (73)$$

Therefore, we found that the recovery map is also a stochastic Pauli map, albeit with some negative coefficients, which render it non-CP. Finally, we sum the modulus of the expansion's coefficients to compute the cost of implementation of $\mathcal{R}$, according to the quasiprobability decomposition in Eq. (72). After some algebra, we obtain

$$\gamma_{\mathcal{R}} = \frac{1}{|p_0 - p_1 - p_2 + p_3|}. \quad (74)$$

Similarly to the previous example, this cost is optimal [22], meaning that no non-CPTP operation with the associated recovery property can be simulated by the quasiprobability decomposition method with a smaller sampling overhead. Analogous recovery maps, along with their corresponding sampling costs and optimality properties, can be derived for the other Pauli observables.

The optimal costs presented in Eqs. (62) and (74) correspond to sampling overhead implementing the information recovery map for a specific observable O ($O = X$ for the

preprocessing of the GAD channel and $O = Z$ for the postprocessing of the stochastic Pauli channel). This is the underlying reason why our protocol could be advantageous over implementing the inverse map $\mathcal{E}^{-1}$, which fully recovers any observable O . We also note that these decompositions are optimized over all possible QPDs using CPTP maps of the form in Eq. (39). However, as the optimal decompositions for both cases are given by weighted sums of Pauli operations, the preprocessing and postprocessing protocols could be readily implemented in a wide range of experimental settings.

VII. CONCLUSIONS

In this work, we introduced the concept of QOOT as a dual object of the QSOT, allowing us to jointly describe two observables that are timelike separated, while also capturing their temporal correlations. The main difference compared with QSOT is that the construction of the QOOT is not universally applicable. We established a no-go theorem for constructing QOOT by fully characterizing the condition on the pairs of prior observables and noise maps to properly define a QOOT. We then leveraged QOOTs to systematically define recovery maps via an equation that reflects a certain time-reversal symmetry of the system, establishing a notion of time reversal for generic quantum channels, with respect to a reference observable. By construction, these time-reversal maps act as recovery maps for the reference observable and, although generally non-CP, they can be decomposed into physically realizable maps, enabling their application in error mitigation tasks. We introduced a specific instance of the QOOT, based on the Jordan product, and used it to provide explicit examples of recovery maps for widely employed noise models. For the noise models we studied, our approach achieves optimal performance, measured in terms of sampling overhead, outperforming the traditional probabilistic error cancellation protocol.

This work leaves interesting open questions that are worth exploring in future research. First, it would be valuable to investigate how general the optimality result is and to further understand the underlying reasons for its validity. Additionally, while we demonstrated that the QOOT, as defined here, may not always be constructible when the noise channel is nonunital, it remains an open question whether an alternative framework can be developed to ensure well-defined recovery maps under fully general conditions. It would also be interesting to explore extending the formalism to continuous variable systems and to compare the performance of our error mitigation protocol with that of the PEC approach [29].

ACKNOWLEDGMENTS

G.B. was part of the AppQInfo MSCA ITN, which received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 956071. H.K. was supported by KIAS individual Grant No. CG085301 at the Korea Institute for Advanced Study. This work was also supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (Grant No. RS-2024-00413957). M.S.K. acknowledges support from UK

EPSRC No. EP/W032643/1, No. EP/Y004752/1, and No. EP/Z53318X/1.

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

APPENDIX A: USEFUL IDENTITIES

In this Appendix, we prove that for a CPTP map $\mathcal{E}$ it holds true that $\tau(D[\mathcal{E}^\dagger]) = D[\mathcal{E}]$:

$$\begin{aligned} D[\mathcal{E}] &= \sum_{ij} |i\rangle\langle j| \otimes \mathcal{E}(|j\rangle\langle i|) = \sum_{ij\ell} |i\rangle\langle j| \otimes K_\ell |j\rangle\langle i| K_\ell^\dagger \\ &= \sum_{ij\ell nm} |i\rangle\langle j| \otimes |n\rangle\langle n| K_\ell |j\rangle\langle i| K_\ell^\dagger |m\rangle\langle m| \\ &= \sum_{ij\ell nm} |i\rangle\langle i| K_\ell^\dagger |m\rangle\langle n| K_\ell |j\rangle\langle j| \otimes |n\rangle\langle m| \\ &= \sum_{mnl} K_\ell^\dagger |m\rangle\langle n| K_\ell \otimes |n\rangle\langle m| \\ &= \sum_{nm} \mathcal{E}^\dagger(|m\rangle\langle n|) \otimes |n\rangle\langle m|, \end{aligned} \quad (\text{A1})$$

where we have used the Kraus decomposition of $\mathcal{E}$, namely, $\mathcal{E}(\rho) = \sum_\ell K_\ell \rho K_\ell^\dagger$. Hence, we obtain $\tau(D[\mathcal{E}^\dagger]) = (D[\mathcal{E}])^\dagger = D[\mathcal{E}]$, where the last equality follows from $\mathcal{E}$ being CP.

APPENDIX B: PREPROCESSING PROTOCOL RECOVERY MAP

In this Appendix, we obtain the preprocessing map by solving the following equation:

$$\{\mathcal{I} \otimes O, D[\mathcal{E}]\} = \{\mathcal{I} \otimes \mathcal{E}^\dagger(O), D[\mathcal{P}^\dagger]\}. \quad (\text{B1})$$

The reader may use Fig. 1 as a reference for the spaces that maps and operators act upon. Let us consider the eigendecomposition of the noisy reference observable, i.e., $\mathcal{E}^\dagger(O) = \sum_k q_k |\omega_k\rangle\langle\omega_k|$, and project both sides of Eq. (B1) onto the joint Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ basis $\{| \omega_i \rangle \otimes | \omega_j \rangle\} \equiv \{| \omega_i \omega_j \rangle\}$. In particular, the left-hand side reads

$$\begin{aligned} &\langle \omega_i | \langle \omega_k | \{\mathcal{I} \otimes O, D[\mathcal{E}]\} | \omega_j \rangle | \omega_\ell \rangle \\ &= \langle \omega_i | \langle \omega_k | (\mathcal{I} \otimes O) D[\mathcal{E}] | \omega_j \rangle | \omega_\ell \rangle \\ &\quad + \langle \omega_i | \langle \omega_k | D[\mathcal{E}] (\mathcal{I} \otimes O) | \omega_j \rangle | \omega_\ell \rangle \\ &= \langle \omega_i | \langle \omega_k | \{O, \mathcal{E}(|\omega_\ell\rangle\langle\omega_\ell|)\} | \omega_j \rangle, \end{aligned} \quad (\text{B2})$$

where we have used the identity $\langle \omega_k | D[\mathcal{E}] | \omega_\ell \rangle = \mathcal{E}(|\omega_\ell\rangle\langle\omega_k|)$. On the other hand, the right-hand side reads

$$\begin{aligned} &\langle \omega_i | \langle \omega_k | \{\mathcal{I} \otimes \mathcal{E}^\dagger(O), D[\mathcal{P}^\dagger]\} | \omega_j \rangle | \omega_\ell \rangle \\ &= \langle \omega_i | \langle \omega_k | (\mathcal{I} \otimes \mathcal{E}^\dagger(O)) D[\mathcal{P}^\dagger] | \omega_j \rangle | \omega_\ell \rangle \\ &\quad + \langle \omega_i | \langle \omega_k | D[\mathcal{P}^\dagger] (\mathcal{I} \otimes \mathcal{E}^\dagger(O)) | \omega_j \rangle | \omega_\ell \rangle \\ &= (q_k + q_\ell) \langle \omega_i | \langle \omega_k | D[\mathcal{P}^\dagger] | \omega_j \rangle | \omega_\ell \rangle \\ &= (q_k + q_\ell) \langle \omega_i | \mathcal{P}^\dagger(|\omega_\ell\rangle\langle\omega_k|) | \omega_j \rangle. \end{aligned} \quad (\text{B3})$$

Putting everything together and further assuming that $q_k + q_\ell \neq 0$ for all k and ℓ , we finally get to

$$\mathcal{P}^\dagger(|\omega_k\rangle\langle\omega_\ell|) = (q_k + q_\ell)^{-1} \{O, \mathcal{E}(|\omega_k\rangle\langle\omega_\ell|)\}. \quad (\text{B4})$$

APPENDIX C: RECOVERY MAP FOR UNITARY QUANTUM CHANNELS

In this Appendix, we explicitly show that, when the quantum channel $\mathcal{E}$ models a unitary evolution, namely, $\mathcal{E}(\rho) = U \rho U^\dagger$ with $UU^\dagger = \mathcal{I}$, then its inverse $\mathcal{E}^{-1}(\rho) = U^\dagger \rho U$ is a recovery map $\mathcal{P}$ for the preprocessing protocol, i.e., it solves

$$\{\mathcal{I} \otimes O, D[\mathcal{E}]\} = \{\mathcal{I} \otimes \mathcal{E}^\dagger(O), D[\mathcal{P}^\dagger]\}. \quad (\text{C1})$$

We showed in Sec. IV A that the solution to Eq. (C1) is given by

$$\mathcal{P}^\dagger(|\omega_k\rangle\langle\omega_\ell|) = (q_k + q_\ell)^{-1} \{O, \mathcal{E}(|\omega_k\rangle\langle\omega_\ell|)\}, \quad (\text{C2})$$

where q_k and $|\omega_k\rangle$ are the eigenvalues and eigenstates of $\mathcal{E}^\dagger(O)$, respectively. The noisy reference observable reads

$$\mathcal{E}^\dagger(O) = U^\dagger O U = \sum_k q_k U^\dagger |\lambda_k\rangle\langle\lambda_k| U, \quad (\text{C3})$$

where we have used the eigendecomposition of the target observable, namely, $O = \sum_k q_k |\lambda_k\rangle\langle\lambda_k|$. It thus follows that $|\omega_k\rangle = U^\dagger |\lambda_k\rangle$. We can use all of these facts to rewrite Eq. (C2) as

$$\begin{aligned} \mathcal{P}^\dagger(|\omega_k\rangle\langle\omega_\ell|) &= (q_k + q_\ell)^{-1} \{O, \mathcal{E}(|\omega_k\rangle\langle\omega_\ell|)\} \\ &= (q_k + q_\ell)^{-1} \{\sum_i q_i |\lambda_i\rangle\langle\lambda_i|, U |\omega_k\rangle\langle\omega_\ell| U^\dagger\} \\ &= (q_k + q_\ell)^{-1} \sum_i q_i \{|\lambda_i\rangle\langle\lambda_i|, U U^\dagger \\ &\quad \times |\lambda_k\rangle\langle\lambda_\ell| U U^\dagger\} \\ &= (q_k + q_\ell)^{-1} (q_k |\lambda_k\rangle\langle\lambda_\ell| + q_\ell |\lambda_\ell\rangle\langle\lambda_k|) \\ &= |\lambda_k\rangle\langle\lambda_\ell| = U |\omega_k\rangle\langle\omega_\ell| U^\dagger, \end{aligned} \quad (\text{C4})$$

which implies that $\mathcal{P}^\dagger = \mathcal{E}$, or equivalently that $\mathcal{P} = \mathcal{E}^{-1}$, concluding our proof.

APPENDIX D: POSTPROCESSING PROTOCOL RECOVERY MAP (QUBIT SYSTEMS)

Once again, the reader may use Fig. 2 as a reference for the spaces that maps and operators act upon. In the main text, we showed that a recovery map for a postprocessing protocol may be obtained by solving

$$\{O \otimes \mathcal{I}, D[\mathcal{R}^\dagger]\} = \{\mathcal{R}^\dagger(O) \otimes \mathcal{I}, D[\mathcal{E}]\}. \quad (\text{D1})$$

The difficulty in solving the above equation lies in the fact that, unlike the preprocessing scenario, $\mathcal{R}^\dagger$ appears on both sides of the equation and, more importantly, we do not know anything about $\mathcal{R}^\dagger(O)$, let alone its diagonalization. Instead, we consider the eigendecomposition of the reference observable $O = \sum_i q_i |\omega_i\rangle\langle\omega_i|$ and project both sides of the equation above on the joint Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ basis $\{| \omega_i \rangle \otimes | \omega_j \rangle\} \equiv \{| \omega_i \omega_j \rangle\}$. In particular, the left-hand side

reads

$$\begin{aligned}
 & \langle \omega_k | \langle \omega_i | \{O \otimes \mathcal{I}, D[\mathcal{R}^\dagger]\} | \omega_\ell \rangle | \omega_j \rangle \\
 &= \langle \omega_k | \langle \omega_i | (O \otimes \mathcal{I}) D[\mathcal{R}^\dagger] | \omega_\ell \rangle | \omega_j \rangle \\
 &+ \langle \omega_k | \langle \omega_i | D[\mathcal{R}^\dagger] (O \otimes \mathcal{I}) | \omega_\ell \rangle | \omega_j \rangle \\
 &= (q_k + q_\ell) \langle \omega_k | \langle \omega_i | D[\mathcal{R}^\dagger] | \omega_\ell \rangle | \omega_j \rangle \\
 &= (q_k + q_\ell) \langle \omega_i | \mathcal{R}^\dagger(|\omega_\ell\rangle\langle\omega_k|) | \omega_j \rangle. \quad (\text{D2})
 \end{aligned}$$

With similar techniques, we compute the right-and side, namely,

$$\begin{aligned}
 & \langle \omega_k | \langle \omega_i | \{\mathcal{R}^\dagger(O) \otimes \mathcal{I}, D[\mathcal{E}]\} | \omega_\ell \rangle | \omega_j \rangle \\
 &= \langle \omega_i | \{\mathcal{R}^\dagger(O), \mathcal{E}(|\omega_\ell\rangle\langle\omega_k|)\} | \omega_j \rangle. \quad (\text{D3})
 \end{aligned}$$

Putting everything together, we obtain

$$(q_k + q_\ell) \mathcal{R}^\dagger(|\omega_\ell\rangle\langle\omega_k|) = \{\mathcal{R}^\dagger(O), \mathcal{E}(|\omega_\ell\rangle\langle\omega_k|)\}. \quad (\text{D4})$$

Notice that, while we do not know what $\mathcal{R}^\dagger(O)$ is, we may conveniently express it as $\mathcal{R}^\dagger(O) = \sum_i q_i \mathcal{R}^\dagger(|\omega_i\rangle\langle\omega_i|)$, by linearity. Hence, we can rewrite Eq. (D4) as

$$(q_k + q_\ell) \mathcal{R}^\dagger(|\omega_\ell\rangle\langle\omega_k|) = \sum_i q_i \{\mathcal{R}^\dagger(|\omega_i\rangle\langle\omega_i|), \mathcal{E}(|\omega_\ell\rangle\langle\omega_k|)\}. \quad (\text{D5})$$

The equation above can be recast as

$$X_{k\ell} = \sum_i \{X_{ii}, A_{k\ell}\}, \quad (\text{D6})$$

where

$$X_{k\ell} = (q_k + q_\ell) \mathcal{R}^\dagger(|\omega_k\rangle\langle\omega_\ell|), \quad (\text{D7})$$

$$A_{k\ell} = \frac{1}{2} \mathcal{E}(|\omega_k\rangle\langle\omega_\ell|). \quad (\text{D8})$$

All that is left now is to solve Eq. (D6). For a qubit system, we may decompose all operators appearing in the equation in the Pauli basis, namely,

$$X_{k\ell} = \sum_{j=0}^3 x_j^{(k\ell)} \sigma_j, \quad A_{k\ell} = \sum_{j=0}^3 a_j^{(k\ell)} \sigma_j, \quad (\text{D9})$$

where $\{\sigma_0, \sigma_1, \sigma_2, \sigma_3\} \equiv \{\mathcal{I}, X, Y, Z\}$ denote the Pauli operators. We then substitute the above expression into Eq. (D6), yielding

$$\sum_j x_j^{(k\ell)} \sigma_j = \sum_{imn} x_m^{(ii)} a_n^{(k\ell)} \{\sigma_m, \sigma_n\}. \quad (\text{D10})$$

The anticommutator of two Pauli matrices is equal to $\{\sigma_m, \sigma_n\} = 2\mathcal{I}$ for every $m = n$, $\{\sigma_0, \sigma_n\} = 2\sigma_n$, and is zero otherwise. The right-hand side of Eq. (D10) then reads

$$\begin{aligned}
 & \sum_{imn} x_m^{(ii)} a_n^{(k\ell)} \{\sigma_m, \sigma_n\} \\
 &= \sum_i \left[\sum_n x_n^{(ii)} a_n^{(k\ell)} 2\mathcal{I} + \sum_{n \neq 0} (x_0^{(ii)} a_n^{(k\ell)} + x_n^{(ii)} a_0^{(k\ell)}) 2\sigma_n \right]. \quad (\text{D11})
 \end{aligned}$$

Putting everything together, we obtain

$$\begin{aligned}
 \sum_j x_j^{(k\ell)} \sigma_j &= \sum_i \sum_n x_n^{(ii)} a_n^{(k\ell)} 2\mathcal{I} \\
 &+ \sum_i \sum_{n \neq 0} (x_0^{(ii)} a_n^{(k\ell)} + x_n^{(ii)} a_0^{(k\ell)}) 2\sigma_n, \quad (\text{D12})
 \end{aligned}$$

which leads to the following set of linear equations for the (eight) coefficients $x_j^{(k\ell)}$:

$$x_0^{(k\ell)} = 2 \sum_{i,j} x_j^{(ii)} a_j^{(k\ell)}, \quad (\text{D13})$$

$$x_j^{(k\ell)} = 2 \sum_i (x_0^{(ii)} a_j^{(k\ell)} + x_j^{(ii)} a_0^{(k\ell)}), \quad j \in \{1, 2, 3\}. \quad (\text{D14})$$

APPENDIX E: POSTPROCESSING PROTOCOL RECOVERY MAP (d -DIMENSIONAL SYSTEM)

We now want to generalize the expressions and the system of linear equations obtained in the previous section for a generic d -dimensional system. Once again, the idea is to start from $X_{k\ell} = \sum_i \{X_{ii}, A_{k\ell}\}$ and decompose the operators appearing in the equation onto a suitable operator basis. To exploit the structure of the equation, we need a basis that has simple anticommutation properties. For n -qubit systems, we might be tempted to use the Pauli basis $\sigma_{j_1} \otimes \cdots \otimes \sigma_{j_n}$ to decompose our operators; however, it becomes tricky to concisely express the anticommutator of two basis elements. Hence, we resort to the basis constructed starting from the so-called *clock* and *shift* operators. In dimension $d = 2$, these are simply the Pauli operators σ_1 and σ_3 , and they satisfy $\sigma_1^2 = \sigma_3^2 = \mathcal{I}$ and $\sigma_1 \sigma_3 = -\sigma_3 \sigma_1 = e^{i\pi} \sigma_3 \sigma_1$. In arbitrary dimension d ($d = 2^n$ for n -qubit systems), these operators generalize to

$$\begin{aligned}
 \Sigma_1 &= \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}, \\
 \Sigma_3 &= \text{diag}(1, \omega, \omega^2, \dots, \omega^{d-1}), \quad (\text{E1})
 \end{aligned}$$

where $\omega = e^{\frac{2\pi i}{d}}$. These matrices are unitary, traceless but not Hermitian for $d > 2$. They satisfy $\Sigma_1^d = \Sigma_3^d = \mathcal{I}$ as well as $\Sigma_3 \Sigma_1 = \omega \Sigma_1 \Sigma_3$ and, most importantly, they may be used to define the following complete set of d^2 $d \times d$ matrices, namely,

$$\sigma_{j,k} = \Sigma_1^k \Sigma_3^j, \quad (\text{E2})$$

with $j, k \in \{1, \dots, d\}$. By repeatedly using the defining property $\Sigma_3 \Sigma_1 = \omega \Sigma_1 \Sigma_3$, one can also show that $\Sigma_3^j \Sigma_1^k = \omega^{jk} \Sigma_1^k \Sigma_3^j$. It follows that

$$\{\sigma_{j,k}, \sigma_{l,m}\} = (\omega^{kl} + \omega^{mj}) \sigma_{j+l, k+m}. \quad (\text{E3})$$

As a result, for a d -dimensional system we may decompose the operators $X_{k\ell}$ and $A_{k\ell}$ as follows:

$$X_{k\ell} = \sum_{l,m=1}^d x_{lm}^{(k\ell)} \sigma_{l,m}, \quad A_{k\ell} = \sum_{l,m=1}^d a_{lm}^{(k\ell)} \sigma_{l,m}. \quad (\text{E4})$$

Substituting these expressions into $X_{k\ell} = \sum_i \{X_{ii}, A_{k\ell}\}$ and using the anticommutation relation Eq. (E3) leads to

$$\begin{aligned} \sum_{lm} x_{lm}^{(k\ell)} \sigma_{l,m} &= \sum_{istpn} x_{st}^{(ii)} a_{pn}^{(k\ell)} \{\sigma_{s,t}, \sigma_{p,n}\} \\ &= \sum_{istpn} x_{st}^{(ii)} a_{pn}^{(k\ell)} (\omega^{tp} + \omega^{sn}) \sigma_{s+p,t+n}. \end{aligned} \quad (\text{E5})$$

Consequently, we obtain the following linear equation for the coefficients $x_{lm}^{(k\ell)}$:

$$x_{lm}^{(k\ell)} = \sum_i \sum_{s+p=l} \sum_{t+n=m} x_{st}^{(ii)} a_{pn}^{(k\ell)} (\omega^{tp} + \omega^{sn}). \quad (\text{E6})$$

APPENDIX F: NECESSARY AND SUFFICIENT CONDITION FOR EQ. (30)

Here, we prove that the necessary and sufficient condition $\{O, \mathcal{R}(\mathcal{I})\} = 2O$ for the Jordan product QOOT $\mathcal{R}^\dagger \star O = \frac{1}{2} \{O \otimes \mathcal{I}, D[\mathcal{R}^\dagger]\}$ to satisfy its defining property, is equivalent to imposing $\text{Tr}[X_{k\ell}] = q_k \delta_{k\ell}$, where $X_{k\ell} = (q_k + q_\ell) \mathcal{R}^\dagger(|\omega_k\rangle\langle\omega_\ell|)$. We remind the reader that q_k and $|\omega_k\rangle$ are

the eigenvalues and eigenstates of the reference observable O , respectively:

$$\begin{aligned} \frac{1}{2} \{O, \mathcal{R}(\mathcal{I})\} &= \frac{1}{2} \text{Tr}_B[\{O \otimes \mathcal{I}, D[\mathcal{R}^\dagger]\}] \\ &= \frac{1}{2} \sum_{k\ell} \{O, |\omega_k\rangle\langle\omega_\ell|\} \text{Tr}[\mathcal{R}^\dagger(|\omega_\ell\rangle\langle\omega_k|)] \\ &= \frac{1}{2} \sum_{k\ell} (q_k + q_\ell) |\omega_k\rangle\langle\omega_\ell| \text{Tr}[\mathcal{R}^\dagger(|\omega_\ell\rangle\langle\omega_k|)] \\ &= \frac{1}{2} \sum_{k\ell} \text{Tr}[(q_k + q_\ell) \mathcal{R}^\dagger(|\omega_\ell\rangle\langle\omega_k|)] |\omega_k\rangle\langle\omega_\ell| \\ &= \frac{1}{2} \sum_{k\ell} \text{Tr}[X_{k\ell}] |\omega_k\rangle\langle\omega_\ell|, \end{aligned} \quad (\text{F1})$$

where in the second equality we have exploited the definition of the Jamiołkowski state [Eq. (12)], using $\{|\omega_k\rangle\}$ as the basis of choice. The chain of identities above implies that imposing $\frac{1}{2} \{O, \mathcal{R}(\mathcal{I})\} = O$ is equivalent to imposing $\frac{1}{2} \sum_{k\ell} \text{Tr}[X_{k\ell}] |\omega_k\rangle\langle\omega_\ell| = \sum_k q_k |\omega_k\rangle\langle\omega_k|$. Finally, this last constraint is clearly satisfied if and only if $\text{Tr}[X_{k\ell}] = q_k \delta_{k\ell}$.

-
- [1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2010).
 - [2] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, Cambridge, 2013).
 - [3] J. Fullwood and A. J. Parzygnat, On quantum states over time, *Proc. R. Soc. A* **478**, 20220104 (2022).
 - [4] A. J. Parzygnat and F. Buscemi, Axioms for retrodiction: Achieving time-reversal symmetry with *a priori*, *Quantum* **7**, 1013 (2023).
 - [5] A. J. Parzygnat and J. Fullwood, From time-reversal symmetry to quantum Bayes' rules, *PRX Quantum* **4**, 020334 (2023).
 - [6] D. Petz, Sufficient subalgebras and the relative entropy of states of a von Neumann algebra, *Commun. Math. Phys.* **105**, 123 (1986).
 - [7] H. Kwon, R. Mukherjee, and M. S. Kim, Reversing Lindblad dynamics via continuous Petz recovery map, *Phys. Rev. Lett.* **128**, 020403 (2022).
 - [8] H. Barnum and E. Knill, Reversing quantum dynamics with near-optimal quantum and classical fidelity, *J. Math. Phys.* **43**, 2097 (2002).
 - [9] M. M. Wilde, Recoverability in quantum information theory, *Proc. R. Soc. A* **471**, 20150338 (2015).
 - [10] K. Temme, S. Bravyi, and J. M. Gambetta, Error mitigation for short-depth quantum circuits, *Phys. Rev. Lett.* **119**, 180509 (2017).
 - [11] M. S. Leifer and R. W. Spekkens, Towards a formulation of quantum theory as a causally neutral theory of Bayesian inference, *Phys. Rev. A* **88**, 052130 (2013).
 - [12] D. Horsman, C. Heunen, M. F. Pusey, J. Barrett, and R. W. Spekkens, Can a quantum state over time resemble a quantum state at a single time? *Proc. R. Soc. A* **473**, 20170395 (2017).
 - [13] M. Ohya, Note on quantum probability, *Lett. Nuovo Cimento* **38**, 402 (1983).
 - [14] M. Ohya, On compound state and mutual information in quantum information theory (corresp.), *IEEE Trans. Inf. Theory* **29**, 770 (1983).
 - [15] M. Asorey, A. Kossakowski, G. Marmo, and E. G. Sudarshan, Relations between quantum maps and quantum states, *Open Syst. Inf. Dyn.* **12**, 319 (2005).
 - [16] S. H. Lie and N. H. Y. Ng, Quantum state over time is unique, *Phys. Rev. Res.* **6**, 033144 (2024).
 - [17] J. F. Fitzsimons, J. A. Jones, and V. Vedral, Quantum correlations which imply causation, *Sci. Rep.* **5**, 18281 (2015).
 - [18] M. S. Leifer, Quantum dynamics as an analog of conditional probability, *Phys. Rev. A* **74**, 042310 (2006).
 - [19] F. Buscemi, M. Dall'Arno, M. Ozawa, and V. Vedral, Direct observation of any two-point quantum correlation function, *arXiv:1312.4240*.
 - [20] J. Jiang, K. Wang, and X. Wang, Physical implementability of linear maps and its application in error mitigation, *Quantum* **5**, 600 (2021).
 - [21] S. Endo, S. C. Benjamin, and Y. Li, Practical quantum error mitigation for near-future applications, *Phys. Rev. X* **8**, 031027 (2018).
 - [22] X. Zhao, B. Zhao, Z. Xia, and X. Wang, Information recoverability of noisy quantum states, *Quantum* **7**, 978 (2023).
 - [23] A. Jamiołkowski, Linear transformations which preserve trace and positive semidefiniteness of operators, *Rep. Math. Phys.* **3**, 275 (1972).
 - [24] W. Hoeffding, Probability inequalities for sums of bounded random variables, *J. Am. Stat. Assoc.* **58**, 13 (1963).

- [25] X. Zhao, L. Zhang, B. Zhao, and X. Wang, Power of quantum measurement in simulating unphysical operations, [Phys. Rev. Res. **7**, 013334 \(2025\)](#).
- [26] Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O'Brien, Quantum error mitigation, [Rev. Mod. Phys. **95**, 045005 \(2023\)](#).
- [27] C. Dankert, R. Cleve, J. Emerson, and E. Livine, Exact and approximate unitary 2-designs and their application to fidelity estimation, [Phys. Rev. A **80**, 012304 \(2009\)](#).
- [28] S. Khatri, K. Sharma, and M. M. Wilde, Information-theoretic aspects of the generalized amplitude-damping channel, [Phys. Rev. A **102**, 012401 \(2020\)](#).
- [29] A. Taylor, G. Bressanini, H. Kwon, and M. S. Kim, Quantum error cancellation in photonic systems: Undoing photon losses, [Phys. Rev. A **110**, 022622 \(2024\)](#).