Evaluating the advantage of adaptive strategies for quantum channel distinguishability

Vishal Katariya[1] and Mark M. Wilde[1,2]

[1]*Hearne Institute for Theoretical Physics, Department of Physics and Astronomy, and Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana 70803, USA*

[2]*Stanford Institute for Theoretical Physics, Stanford University, Stanford, California 94305, USA*

Recently, the resource theory of asymmetric distinguishability for quantum strategies was introduced by Wang *et al.* [Phys. Rev. Research **1**, 033169 (2019).]. The fundamental objects in the resource theory are pairs of quantum strategies, which are generalizations of quantum channels that provide a framework to describe an arbitrary quantum interaction. In the present paper we provide semidefinite program characterizations of the one-shot operational quantities in this resource theory. We then apply these semidefinite programs to study the advantage conferred by adaptive strategies in discrimination and distinguishability distillation of generalized amplitude damping channels. We find that there are significant gaps between what can be accomplished with an adaptive strategy versus a nonadaptive strategy.

I. INTRODUCTION

In quantum information theory, the tasks of quantum state and channel discrimination have been studied in a considerable amount of detail; see Refs. [1–5] and [6–11], respectively. Given the central importance of distinguishing quantum states or channels, it is reasonable to study distinguishability itself in the context of a resource theory [12–14], i.e., to use resource-theoretic tools to quantify distinguishability, and to use these tools to study the tasks of distilling distinguishability from a pair of objects, diluting canonical units of distinguishability to a desired pair, and transforming one pair of entities to another pair.

References [12–14] developed some basic tools and a framework for the resource theory of asymmetric distinguishability. In some sense, the resource theory of asymmetric distinguishability can be thought of as a "meta"-resource theory. The basic objects in this resource theory come in pairs, and their worth is decided by the distinguishability of the entities in a pair. This resource theory is also unique in the sense that all physical operations acting on each element of the pair are free. A variety of resource theories can be thought of as being derived from the resource theory of asymmetric distinguishability by setting specific restrictions on the states or channels allowed for free [13].

The most general discrimination task in quantum information theory is not that of discriminating channels but that of distinguishing what are known as quantum strategies [15–17], also known as quantum combs, memory channels, or higher-order quantum maps [18–20]. A quantum strategy completely represents the actions of an agent in a multiround interaction with another party and forms the next rung in the hierarchical ladder that begins with quantum states and channels. A key insight of Ref. [20] is that the hierarchy consisting of states, channels, superchannels, etc., ends with quantum

strategies. That is, all so-called "higher-order" dynamics can be described as quantum strategies. Given this importance of quantum strategies and the flexibility and power offered by the resource theory of asymmetric distinguishability, it is worthwhile to continue the study of it for quantum strategies, as initiated in Ref. [14].

In this paper we provide several contributions to the resource theory of asymmetric distinguishability [13,14]. One of our main contributions is a semidefinite programming (SDP) characterization of two crucial quantities in this resource theory: the one-shot distillable distinguishability and the one-shot distinguishability cost of quantum strategies, which characterize the resource theory's distillation and dilution tasks, respectively. To do so, we build upon the previous SDP characterizations of the quantum strategy distance [17,19], which provides a distance measure between strategies.

The other main contribution of this paper is to apply these SDPs to study particular examples of channel distinguishability tasks. As indicated in Ref. [14], distinguishability distillation is closely linked to asymmetric quantum channel discrimination. In quantum channel discrimination, one can employ either parallel or adaptive strategies. By definition, adaptive strategies are no less powerful than parallel ones. It is known that in the asymptotic limit, adaptive strategies confer no advantage over parallel ones in asymmetric channel discrimination [21–23]. This leaves open the question of whether adaptive strategies can help in channel discrimination when a finite number of channel uses are allowed. Our SDP formulations help us compute and study this gap. As an example, we consider distinguishability tasks involving generalized amplitude damping channels (GADCs) and show that adaptive strategies offer a significant advantage over parallel ones with respect to various distinguishability metrics of interest, thus extending prior work on this topic from Ref. [24].
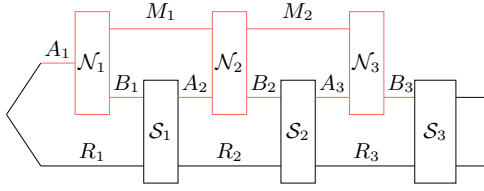
FIG. 1. A three-turn strategy $\mathcal{N}^{(3)}$ interacts with a three-turn co-strategy $\mathcal{S}^{(3)}$. In red is the entirety of the three-turn strategy $\mathcal{N}^{(3)}$, which consists of three quantum channels $\mathcal{N}_1$ through $\mathcal{N}_3$, connected to each other via memory systems. The co-strategy $\mathcal{S}^3$ consists of an initial state on systems A_1R_1 as well, the quantum channels $\mathcal{S}_1$, $\mathcal{S}_2$, and $\mathcal{S}_3$, and memory systems R_1, R_2, and R_3.

II. QUANTUM STRATEGIES

The idea of quantum strategies, combs, or higher-order maps goes back over a decade [15,20]. A quantum strategy generalizes a quantum channel in that it allows for sequential interactions over multiple rounds. Consider that there are two parties, Alice and Bob. Alice's n-turn quantum strategy describes her actions in an n-round interaction with Bob. In such a scenario, Bob's n-round interaction is described by a suitable quantum co-strategy. In other words, the interaction of Alice's n-turn quantum strategy with another suitable n-turn strategy (belonging to Bob) captures all possible interactive behavior that takes place over n rounds between them. Reference [20] introduced the term "quantum comb," which refers to the same physical object as a quantum strategy.

An n-turn quantum strategy $\mathcal{N}^{(n)}$, with $n \geqslant 1$, input systems A_1 through A_n, and output systems B_1 through B_n, consists of the following: (a) memory systems M_1 through M_{n-1}, and (b) quantum channels $\mathcal{N}^1_{A_1 \to M_1 B_1}$, $\mathcal{N}^2_{M_1 A_2 \to M_2 B_2}$, ..., $\mathcal{N}^{n-1}_{M_{n-2} A_{n-1} \to M_{n-1} B_{n-1}}$, and $\mathcal{N}^n_{M_{n-1} A_n \to B_n}$. The definition above allows for any of the input, output, or memory systems to be trivial, which means that state preparation and measurements can be captured in the framework of quantum strategies. For the sake of brevity, we use the notation A^n to denote systems A_1 through A_n. Figure 1 depicts a three-turn strategy interacting with a three-turn co-strategy.

A superchannel is a physical operation that converts one quantum channel to another [18,25]. It is a particular type of quantum strategy. Reference [18] made the important observation that a superchannel can be equivalently represented as a bipartite channel, along with a causality constraint that defines the causal order of inputs and outputs. Reference [20]'s observation that quantum combs are all that are needed to describe higher-order quantum dynamics ties in neatly with, and generalizes, the superchannel-bipartite channel isomorphism. A superchannel can be cast as a bipartite channel, and likewise, an object that transforms superchannels to superchannels (which is a quantum strategy) is itself a multipartite superchannel, which by the previously stated isomorphism is a multipartite channel [20]. Therefore there is a "collapse" of the hierarchy that proves to be important, which implies that all higher-order quantum dynamics can be studied using the framework of quantum strategies [20].

Another isomorphism that is crucial in quantum information is the Choi isomorphism. It too establishes an equivalence between two different classes of objects—a single-party quantum channel can be equivalently represented by a bipartite quantum state. Putting the pieces together, we see that one can define a Choi state, or a Choi operator, not only for quantum channels, but also in general for quantum strategies. This isomorphism enables us to apply the tools developed in the resource theory of asymmetric distinguishability for states and channels to superchannels and, more generally, to quantum strategies. This was identified and studied in Ref. [14], and here we elaborate in much more detail on these points.

In the remainder of this section we establish some preliminaries regarding the quantum strategies formalism, and we also provide a semidefinite program for the quantum strategy distance between two n-round strategies that is slightly different from that presented in Ref. [17].

A. Choi operator and causality constraints

To establish the Choi operator for a quantum strategy, we recall that a superchannel $\Theta_{(A_1 \to B_1) \to (A_2 \to B_2)}$ transforming $\mathcal{N}_{A_1 \to B_1}$ to $\mathcal{K}_{A_2 \to B_2}$ is in one-to-one correspondence with a bipartite channel $\mathcal{L}_{A_2 B_1 \to A_1 B_2}$ that has a certain no-signaling constraint [18,25]. The superchannel $\Theta_{(A_1 \to B_1) \to (A_2 \to B_2)}$ can be implemented via preprocessing and postprocessing channels $\mathcal{E}_{A_2 \to A_1 M}$ and $\mathcal{D}_{B_1 M \to B_2}$ that share a memory system M. The Choi operator of the superchannel $\Theta_{(A_1 \to B_1) \to (A_2 \to B_2)}$, given by $\Gamma^\Theta_{A_1 A_2 B_1 B_2}$, is identified with the Choi operator of the corresponding bipartite channel

$$\mathcal{L}_{A_2 B_1 \to A_1 B_2} := \mathcal{D}_{B_1 M \to B_2} \circ \mathcal{E}_{A_2 \to A_1 M}, \tag{1}$$

along with a causality constraint that ensures no backward signaling in time; i.e., the A systems can signal to the B systems, but not vice versa. This is mathematically represented as

$$\Gamma^\Theta_{A_1 A_2 B_1} = \Gamma^\Theta_{A_1 A_2} \otimes \pi_{B_1}, \tag{2}$$

where π_{B_1} is the maximally mixed state. This reasoning can be extended to quantum strategies.

A general n-turn quantum strategy $\mathcal{N}^{(n)} : A^n \to B^n$ is uniquely associated to its Choi operator $\Gamma^{\mathcal{N}^{(n)}}_{A^n B^n}$ via [15]

$$\Gamma^{\mathcal{N}^{(n)}}_{A^n B^n} := \mathcal{N}^{(n)}_{A'^n \to B^n}(\Gamma_{A'_1 A_1} \otimes \Gamma_{A'_2 A_2} \otimes \cdots \otimes \Gamma_{A'_n A_n}), \tag{3}$$

where $\Gamma_{A'A} \equiv |\Gamma\rangle\langle\Gamma|_{A'A}$, and $|\Gamma\rangle_{A'A} = \sum_i |i\rangle_{A'}|i\rangle_A$ is the unnormalized maximally entangled vector on systems $A'A$. The constraints on the Choi operator $\Gamma^{\mathcal{N}^{(n)}}_{A^n B^n}$ are that

$$\Gamma^{\mathcal{N}^{(n)}}_{A^n B^n} \geqslant 0, \tag{4}$$

and that there exist n positive semidefinite operators $N_{[1]}$, $N_{[2]}$, ..., $N_{[n]}$, with $N_{[i]}$ acting on systems $A^i B^i$ for $1 \leqslant i \leqslant n$, such that

$$N_{[n]} = \Gamma^{\mathcal{N}^{(n)}}_{A^n B^n},$$

$$\mathrm{Tr}_{B_n}\left[N_{[n]}\right] = N_{[n-1]} \otimes I_{A_n},$$

$$\mathrm{Tr}_{B_{n-1}}\left[N_{[n-1]}\right] = N_{[n-2]} \otimes I_{A_{n-1}},$$

$$\vdots \tag{5}$$

$$\mathrm{Tr}_{B_2}\left[N_{[2]}\right] = N_{[1]} \otimes I_{A_2},$$

$$\mathrm{Tr}_{B_1}\left[N_{[1]}\right] = I_{A_1}.$$

These latter constraints (5) are causality constraints that arise due to the flow of information in the strategy. They generalize the single causality constraint imposed on the Choi operator of a superchannel. Conversely, if an operator $\Gamma_{A^n B^n}^{\mathcal{N}^{(n)}}$ satisfies the above constraints, then there is a quantum strategy associated to it [15].

B. Link product

How do we "connect" or compose two quantum strategies? To this end, the notion of link product was introduced to denote the composition, or interaction, of two quantum strategies [20]. Two quantum strategies are composed by connecting the appropriate input and output systems, with an example being given in Fig. 1. Suppose that n-turn strategy $\mathcal{N}^{(n)}$ takes systems A^n to B^n and m-turn strategy $\mathcal{S}^{(m)}$ takes systems C^m to D^m. The Choi operator of the composition $\mathcal{N}^{(n)} \circ \mathcal{S}^{(m)}$ is given by $\Gamma^{\mathcal{N}^{(n)}} * \Gamma^{\mathcal{S}^{(m)}}$, defined in (10). Here, the nomenclature "comb" shines, as we connect the two strategies as if they were interlocking pieces, making sure to connect the appropriate input and output ports of the first and second strategy, respectively.

Qualitatively, the link product connects and "collapses" matching input and output systems of the two strategies. The composition $\mathcal{N}^{(n)} \circ \mathcal{S}^{(m)}$ is another strategy that takes systems

$$(A^n \setminus D^m)(C^m \setminus B^n) \to (D^m \setminus A^n)(B^n \setminus C^m). \quad (6)$$

The matching systems in this case are $A^n \cap D^m$ and $B^n \cap C^m$. To maintain brevity, we define

$$N \cap S := (A^n \cap D^m) \cup (B^n \cap C^m), \quad (7)$$

as well as

$$N \setminus S := (A^n B^n) \setminus (C^m D^m) \quad (8)$$

and

$$S \setminus N := (C^m D^m) \setminus (A^n B^n). \quad (9)$$

The Choi operator of the composition $\mathcal{N}^{(n)} \circ \mathcal{S}^{(n)}$ is given by the link product of strategy Choi operators $\Gamma_{A^n B^n}^{\mathcal{N}^{(n)}}$ and $\Gamma_{C^m D^m}^{\mathcal{S}^{(m)}}$, and is defined as follows:

$$\Gamma_{A^n B^n}^{\mathcal{N}^{(n)}} * \Gamma_{C^m D^m}^{\mathcal{S}^{(m)}} := \mathrm{Tr}_{N \cap S}[(I_{N \setminus S} \otimes (\Gamma^{\mathcal{S}})^{T_{N \cap S}})(\Gamma^{\mathcal{N}} \otimes I_{S \setminus N})], \quad (10)$$

where the notation $T_{N \cap S}$ refers to taking the partial transpose on systems $N \cap S$.

C. Telling two strategies apart

It is natural to introduce a notion of distance, or distinguishability, between two strategies. In this vein, the quantities quantum strategy distance [17,19,20], the strategy fidelity [26], and the strategy max-relative entropy [27] were previously defined. These are generalized by the generalized strategy divergence of Ref. [14].

Given two n-turn strategies with the same input and output systems, the most general discrimination strategy is defined analogously to that in channel discrimination; instead of passing a common state to two channels, one interacts a common n-turn co-strategy with the unknown strategy to obtain an output state on which a measurement is performed.

That is, for strategies $\mathcal{N}^{(n)}$ and $\mathcal{M}^{(n)} : A^n \to B^n$, consider an arbitrary n-turn co-strategy $\mathcal{S}^{(n)} : B^{n-1} \to A^n R_n$. The compositions $\mathcal{N}^{(n)} \circ \mathcal{S}^{(n)}$ and $\mathcal{M}^{(n)} \circ \mathcal{S}^{(n)}$ yield states on $R_n B_n$. The strategy distance between $\mathcal{N}^{(n)}$ and $\mathcal{M}^{(n)}$ is the maximum trace distance between the states on $R_n B_n$ corresponding to strategies $\mathcal{N}^{(n)}$ and $\mathcal{M}^{(n)}$:

$$\left\| \mathcal{N}^{(n)} - \mathcal{M}^{(n)} \right\|_{\diamond n} := \sup_{\mathcal{S}^{(n)}} \left\| \mathcal{N}^{(n)} \circ \mathcal{S}^{(n)} - \mathcal{M}^{(n)} \circ \mathcal{S}^{(n)} \right\|_1. \quad (11)$$

The quantum strategy distance denotes the maximum classical trace distance between the output probability distributions produced by processing both strategies with a common co-strategy. For two arbitrary n-turn strategies, the strategy distance can be computed via a semidefinite program [17], which provides a powerful tool that can be used to study, among other things, the advantage provided by adaptive strategies over parallel ones in quantum channel discrimination, explored in Sec. IV.

In what follows we present an SDP for the *normalized* quantum strategy distance $\frac{1}{2} \| \mathcal{N}^{(n)} - \mathcal{M}^{(n)} \|_{\diamond n}$ of two strategies that is slightly different from that presented previously, in Ref. [17]. This alternate form of the strategy distance is used later to derive SDP characterizations of the distillable distinguishability and the distinguishability cost in Sec. III C.

Proposition 1. The normalized strategy distance $\frac{1}{2} \| \mathcal{N}^{(n)} - \mathcal{M}^{(n)} \|_{\diamond n}$ can be expressed as the following SDP, where $\Gamma_{A^n B^n}^{\mathcal{N}^{(n)}}$ and $\Gamma_{A^n B^n}^{\mathcal{M}^{(n)}}$ are the Choi operators of the strategies $\mathcal{N}_{A^n \to B^n}^{(n)}$ and $\mathcal{M}_{A^n \to B^n}^{(n)}$:

$$\sup_{S, S_{[n]}, \cdots, S_{[1]} \geqslant 0} \left\{ \begin{array}{c} \mathrm{Tr}\left[S(\Gamma^{\mathcal{N}^{(n)}} - \Gamma^{\mathcal{M}^{(n)}}) \right] : \\ S \leqslant S_{[n]} \otimes I_{B_n}, \\ \mathrm{Tr}_{A_n}\left[S_{[n]} \right] = S_{[n-1]} \otimes I_{B_{n-1}}, \\ \vdots \\ \mathrm{Tr}_{A_2}\left[S_{[2]} \right] = S_{[1]} \otimes I_{B_1}, \\ \mathrm{Tr}\left[S_{[1]} \right] = 1 \end{array} \right\}. \quad (12)$$

The dual of the normalized strategy distance is

$$\inf_{\substack{\mu \in \mathbb{R}, Y_n \geqslant 0, \\ Y_1, \ldots, Y_{n-1} \in \mathrm{Herm}}} \left\{ \begin{array}{c} \mu : \\ Y_n \geqslant \Gamma^{\mathcal{N}^{(n)}} - \Gamma^{\mathcal{M}^{(n)}}, \\ Y_{n-1} \otimes I_{A_n} \geqslant \mathrm{Tr}_{B_n}[Y_n], \\ Y_{n-2} \otimes I_{A_{n-1}} \geqslant \mathrm{Tr}_{B_{n-1}}[Y_{n-1}], \\ \vdots \\ Y_1 \otimes I_{A_2} \geqslant \mathrm{Tr}_{B_2}[Y_2], \\ \mu I_{A_1} \geqslant \mathrm{Tr}_{B_1}[Y_1] \end{array} \right\}. \quad (13)$$

Proof. We start by recalling the SDP formulation of the strategy distance from Ref. [17]:

$$\left\| \mathcal{N}^{(n)} - \mathcal{M}^{(n)} \right\|_{\diamond n}$$

$$= \sup_{T_0, T_1 \geqslant 0} \left\{ \begin{array}{c} \operatorname{Tr}[(\Gamma^{\mathcal{N}^{(n)}} - \Gamma^{\mathcal{M}^{(n)}})(T_0 - T_1)] : \\ T_0 + T_1 = T_{[n]} \otimes I_{B_n}, \\ \operatorname{Tr}_{A_n}[T_{[n]}] = T_{[n-1]} \otimes I_{B_{n-1}}, \\ \vdots \\ \operatorname{Tr}_{A_2}[T_{[2]}] = T_{[1]} \otimes I_{B_1}, \\ \operatorname{Tr}[T_{[1]}] = 1 \end{array} \right\}. \quad (14)$$

In the above, $\{T_0, T_1\}$ comprise the Choi operators of an n-round measuring co-strategy, as defined in Ref. [17]. This means that $T = T_0 + T_1$ is the Choi operator of an n-round nonmeasuring co-strategy and obeys the following constraints:

$$T \geqslant 0, \quad (15)$$

$$T = T_{[n]} \otimes I_{B_n}, \quad (16)$$

$$\operatorname{Tr}_{A_n}\left[T_{[n]}\right] = T_{[n-1]} \otimes I_{B_{n-1}}, \quad (17)$$

$$\vdots \quad (18)$$

$$\operatorname{Tr}_{A_2}\left[T_{[2]}\right] = T_{[1]} \otimes I_{B_1}, \quad (19)$$

$$\operatorname{Tr}\left[T_{[1]}\right] = 1. \quad (20)$$

The objective function

$$\operatorname{Tr}[(\Gamma^{\mathcal{N}^{(n)}} - \Gamma^{\mathcal{M}^{(n)}})(T_0 - T_1)] \quad (21)$$

can be rewritten as follows:

$$\operatorname{Tr}[(\Gamma^{\mathcal{N}^{(n)}} - \Gamma^{\mathcal{M}^{(n)}})(T_0 - T_1)]$$

$$= \operatorname{Tr}[(\Gamma^{\mathcal{N}^{(n)}} - \Gamma^{\mathcal{M}^{(n)}})(T_0 - (T - T_0))]$$

$$= 2\operatorname{Tr}[(\Gamma^{\mathcal{N}^{(n)}} - \Gamma^{\mathcal{M}^{(n)}})T_0] - \operatorname{Tr}[(\Gamma^{\mathcal{N}^{(n)}} - \Gamma^{\mathcal{M}^{(n)}})T]$$

$$= 2\operatorname{Tr}[(\Gamma^{\mathcal{N}^{(n)}} - \Gamma^{\mathcal{M}^{(n)}})T_0]. \quad (22)$$

In the above, the first equality arises because $T = T_0 + T_1$. The third equality is due to the fact that for the strategy $\mathcal{N}^{(n)}$, the probabilities of obtaining outcomes corresponding to T_0 and T_1 add to 1, i.e.,

$$\operatorname{Tr}[\Gamma^{\mathcal{N}^{(n)}}(T_0 + T_1)] = \operatorname{Tr}[\Gamma^{\mathcal{N}^{(n)}}T] = 1. \quad (23)$$

The same holds for $\mathcal{M}^{(n)}$ as well, and thus $\operatorname{Tr}[(\Gamma^{\mathcal{N}^{(n)}} - \Gamma^{\mathcal{M}^{(n)}})T] = 0$.

The operator T_0 is the Choi operator of an n-round measuring co-strategy and so it obeys the following constraints:

$$T_0 \geqslant 0, \quad (24)$$

$$T_0 \leqslant T_{[n]} \otimes I_{B_n}, \quad (25)$$

$$\operatorname{Tr}_{A_n}\left[T_{[n]}\right] = T_{[n-1]} \otimes I_{B_{n-1}}, \quad (26)$$

$$\vdots \quad (27)$$

$$\operatorname{Tr}_{A_2}\left[T_{[2]}\right] = T_{[1]} \otimes I_{B_1}, \quad (28)$$

$$\operatorname{Tr}[T_{[1]}] = 1. \quad (29)$$

By combining (22) with the constraints on T_0, we arrive at the desired semidefinite program in (12). Note also that we can

start with (12) and run the whole proof backwards to arrive at (14), setting $T_0 = S$, $T_1 = S_{[n]} \otimes I_{B_n} - S$, and $T = S_{[n]} \otimes I_{B_n}$.

The dual is given by (13), which can be verified by the Lagrange multiplier method and is worked out in Appendix B 2. ∎

III. DISTINGUISHABILITY RESOURCE THEORY

We first recall some aspects of the resource theory of asymmetric distinguishability, work on which was begun in Ref. [12] and continued in Refs. [13,14,28]. The objects in consideration in this resource theory are pairs of like objects. These objects can be probability distributions, quantum states, quantum channels, or most generally, quantum strategies of an equal number of rounds. Any operation on the pair of elements is considered free, justified by the fact that data processing cannot increase the distinguishability of two objects.

The object (ρ, σ), a state box, is an ordered pair of states that is to be understood as an atomic entity: upon being handed a state box, one does not know which state it contains. In this paper we consider ordered pairs of n-turn quantum strategies, which generally are represented by $(\mathcal{N}^{(n)}, \mathcal{M}^{(n)})$.

A. Bits of asymmetric distinguishability

Here we recall the canonical unit of asymmetric distinguishability (AD) [13]. The state box $(|0\rangle\langle 0|, \pi)$ encapsulates one bit of AD, where

$$\pi := \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) \quad (30)$$

is the maximally mixed qubit state. Defining this unit enables us to quantify the amount of resource present in an arbitrary strategy box. As discussed in Ref. [13], the bit of AD represents a pair of experiments in which the null hypothesis corresponds to preparing $|0\rangle\langle 0|$, and the alternative hypothesis corresponds to preparing π. A number m bits of asymmetric distinguishability corresponds to the box $(|0\rangle\langle 0|^{\otimes m}, \pi^{\otimes m})$. Alternatively, the state box $(|0\rangle\langle 0|, \pi_M)$, with

$$\pi_M := \frac{1}{M}|0\rangle\langle 0| + \left(1 - \frac{1}{M}\right)|1\rangle\langle 1|, \quad (31)$$

contains $\log_2 M$ bits of AD.

B. Distillation and dilution of strategy boxes

Given a strategy box $(\mathcal{N}^{(n)}, \mathcal{M}^{(n)})$, we are interested in two questions: (a) how many bits of AD can be distilled from it, and (b) how many bits of AD are required so that one can dilute them to $(\mathcal{N}^{(n)}, \mathcal{M}^{(n)})$? These quantities are crucial to the resource theory of asymmetric distinguishability. The one-shot versions of these tasks are explained below, and we also provide explicit semidefinite programs for them.

We start by discussing distinguishability distillation. The goal of approximate distillation is to transform a strategy box into as many approximate bits of AD as possible. Quantitatively, the one-shot ε-approximate distillable distinguishability of strategy box $(\mathcal{N}^{(n)}, \mathcal{M}^{(n)})$ is given by

$$D_d^\varepsilon(\mathcal{N}^{(n)}, \mathcal{M}^{(n)})$$

$$:= \log_2 \sup_{\mathcal{S}^{(n)}} \{M : \mathcal{N}^{(n)} \circ \mathcal{S}^{(n)} \approx_\varepsilon |0\rangle\langle 0|, \mathcal{M}^{(n)} \circ \mathcal{S}^{(n)} = \pi_M\},$$

where $\mathcal{S}^{(n)}$ is an n-turn co-strategy that interacts with $\mathcal{N}^{(n)}$ and $\mathcal{M}^{(n)}$ to yield a qubit state, and

$$\mathcal{N}^{(n)} \circ \mathcal{S}^{(n)} \approx_\varepsilon |0\rangle\langle 0| \iff \frac{1}{2}\|\mathcal{N}^{(n)} \circ \mathcal{S}^{(n)} - |0\rangle\langle 0|\|_1 \leqslant \varepsilon. \tag{32}$$

The operational quantity for approximate distillation is the smooth strategy min-relative entropy. The smooth strategy min-relative entropy between n-turn strategies $\mathcal{N}^{(n)}$ and $\mathcal{M}^{(n)}$ is defined as follows:

$$D_{\min}^\varepsilon(\mathcal{N}^{(n)}\|\mathcal{M}^{(n)}) := \sup_{\mathcal{S}^{(n)}} D_{\min}^\varepsilon(\mathcal{N}^{(n)} \circ \mathcal{S}^{(n)}\|\mathcal{M}^{(n)} \circ \mathcal{S}^{(n)}),$$

where $\mathcal{N}^{(n)}$ and $\mathcal{M}^{(n)}$ take systems A^n to B^n, and $\mathcal{S}^{(n)}$ is an n-turn co-strategy that takes systems B^{n-1} to $A^n R_n$. The smooth min-relative entropy of states is defined as [29–31]

$$D_{\min}^\varepsilon(\rho\|\sigma) := -\log_2 \inf_{0 \leqslant \Lambda \leqslant I}\{\mathrm{Tr}[\Lambda\sigma] : \mathrm{Tr}[\Lambda\rho] \geqslant 1 - \varepsilon\}.$$

Distinguishability dilution, on the other hand, is the complementary task to distillation. Approximate dilution refers to the task of transforming $(|0\rangle\langle 0|, \pi_M)$ to approximately one copy of $(\mathcal{N}^{(n)}, \mathcal{M}^{(n)})$ with as small M as possible. The one-shot ε-approximate distinguishability cost of the box $(\mathcal{N}^{(n)}, \mathcal{M}^{(n)})$ is given by the following:

$$D_c^\varepsilon(\mathcal{N}^{(n)}, \mathcal{M}^{(n)})$$
$$:= \log_2 \inf_{\mathcal{S}^{(n)}}\{M : \mathcal{S}^{(n)}(|0\rangle\langle 0|) \approx_\varepsilon \mathcal{N}^{(n)}, \mathcal{S}^{(n)}(\pi_M) = \mathcal{M}^{(n)}\}, \tag{33}$$

where

$$\mathcal{S}^{(n)}(|0\rangle\langle 0|) \approx_\varepsilon \mathcal{N}^{(n)} \iff \frac{1}{2}\|\mathcal{S}^{(n)}(|0\rangle\langle 0|) - \mathcal{S}^{(n)}\|_{\diamond n} \leqslant \varepsilon. \tag{34}$$

For the dilution task, the operational quantity is the smooth strategy max-relative entropy of quantum channels and is defined as

$$D_{\max}^\varepsilon(\mathcal{N}^{(n)}\|\mathcal{M}^{(n)}) := \inf_{\widetilde{\mathcal{N}}^{(n)} \approx_\varepsilon \mathcal{N}^{(n)}} D_{\max}(\widetilde{\mathcal{N}}^{(n)}\|\mathcal{M}^{(n)}), \tag{35}$$

where $D_{\max}(\widetilde{\mathcal{N}}^{(n)}\|\mathcal{M}^{(n)})$ is equal to the max-relative entropy for strategies, defined as [27]

$$D_{\max}(\widetilde{\mathcal{N}}^{(n)}\|\mathcal{M}^{(n)}) := D_{\max}(\Gamma^{\widetilde{\mathcal{N}}^{(n)}}\|\Gamma^{\mathcal{M}^{(n)}}), \tag{36}$$

and the max-relative entropy for states is defined as $D_{\max}(\rho\|\sigma) := \inf\{\lambda : \rho \leqslant 2^\lambda \sigma\}$ [32].

We now state a result claimed in Ref. [14]. Its detailed proof is given in Appendix A.

Theorem 2. As seen in Ref. [14], the approximate one-shot distillable distinguishability of the strategy box $(\mathcal{N}^{(n)}, \mathcal{M}^{(n)})$ is equal to the smooth strategy min-relative entropy:

$$D_d^\varepsilon(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) = D_{\min}^\varepsilon(\mathcal{N}^{(n)}\|\mathcal{M}^{(n)}), \tag{37}$$

and the approximate one-shot distinguishability cost is equal to the smooth strategy max-relative entropy:

$$D_c^\varepsilon(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) = D_{\max}^\varepsilon(\mathcal{N}^{(n)}\|\mathcal{M}^{(n)}). \tag{38}$$

C. SDPs for one-shot quantities

This section contains one of our main contributions: explicit semidefinite programs to calculate, for a given strategy box, the approximate distillable distinguishability and approximate distinguishability cost.

Proposition 3. Considering strategies $\mathcal{N}^{(n)}$ and $\mathcal{M}^{(n)}$ to take systems A^n to B^n, the distillable distinguishability is computable via the following semidefinite program:

$$2^{-D_{\min}^\varepsilon(\mathcal{N}^{(n)}\|\mathcal{M}^{(n)})}$$

$$= \inf_{S, S_{[n]}, \ldots, S_{[1]} \geqslant 0}\left\{\begin{array}{c}\mathrm{Tr}[S\,\Gamma^{\mathcal{M}^{(n)}}] : \\ \mathrm{Tr}[S\,\Gamma^{\mathcal{N}^{(n)}}] \geqslant 1 - \varepsilon, \\ S \leqslant S_{[n]} \otimes I_{B_n}, \\ \mathrm{Tr}_{A_n}[S_{[n]}] = S_{[n-1]} \otimes I_{B_{n-1}}, \\ \vdots \\ \mathrm{Tr}_{A_2}[S_{[2]}] = S_{[1]} \otimes I_{B_1}, \\ \mathrm{Tr}[S_{[1]}] = 1\end{array}\right\}, \tag{39}$$

with dual

$$\sup_{\substack{\mu_1, Y_n \geqslant 0, \\ \mu_2 \in \mathbb{R}, \\ Y_1, \ldots, \\ Y_{n-1} \in \mathrm{Herm}}}\left\{\begin{array}{c}(1-\varepsilon)\mu_1 - \mu_2 : \\ Y_n \geqslant \mu_1 \Gamma^{\mathcal{N}^{(n)}} - \Gamma^{\mathcal{M}^{(n)}}, \\ Y_{n-1} \otimes I_{A_n} \geqslant \mathrm{Tr}_{B_n}[Y_n], \\ Y_{n-2} \otimes I_{A_{n-1}} \geqslant \mathrm{Tr}_{B_{n-1}}[Y_{n-1}], \\ \vdots \\ Y_1 \otimes I_{A_2} \geqslant \mathrm{Tr}_{B_2}[Y_2], \\ \mu_2 I_{A_1} \geqslant \mathrm{Tr}_{B_1}[Y_1]\end{array}\right\}. \tag{40}$$

Proof. We have

$$D_d^\varepsilon(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) = D_{\min}^\varepsilon(\mathcal{N}^{(n)}\|\mathcal{M}^{(n)}) \tag{41}$$

$$= \sup_{\mathcal{S}^{(n)}} D_{\min}^\varepsilon(\mathcal{N}^{(n)} \circ \mathcal{S}^{(n)}\|\mathcal{M}^{(n)} \circ \mathcal{S}^{(n)}) \tag{42}$$

and

$$D_{\min}^\varepsilon(\rho\|\sigma)$$
$$= -\log_2 \inf_{\Lambda \geqslant 0}\{\mathrm{Tr}[\Lambda\sigma] : \mathrm{Tr}[\Lambda\rho] \geqslant 1 - \varepsilon, \Lambda \leqslant I\}. \tag{43}$$

We consider $\mathcal{S}^{(n)} : B^{n-1} \to A^n R_n$ to be a co-strategy, so that $\mathcal{N}^{(n)} \circ \mathcal{S}^{(n)}$ is a quantum state on $R_n B_n$. Let $\Lambda_{R_n B_n}$ be a measurement operator such that $\mathrm{Tr}[\Lambda_{R_n B_n}(\mathcal{N}^{(n)} \circ \mathcal{S}^{(n)})]$ is a probability. We now have

$$D_{\min}^\varepsilon(\mathcal{N}^{(n)}\|\mathcal{M}^{(n)}) = -\log_2 \inf_{\mathcal{S}^{(n)}}\left\{\begin{array}{c}\mathrm{Tr}[S^T \Gamma^{\mathcal{M}^{(n)}}] : \\ \mathrm{Tr}[S^T \Gamma^{\mathcal{N}^{(n)}}] \geqslant 1 - \varepsilon\end{array}\right\} \tag{44}$$

such that S is the Choi operator of a valid "sub co-strategy" corresponding to $\mathcal{S}^{(n)}$ and Λ and we have exploited the link product from (10). To write it out explicitly, we use the following constraints on the Choi operator of a sub co-strategy [17, Sec. 2.3]:

$$0 \leqslant S \leqslant S_{[n]} \otimes I_{B_n}, \tag{45}$$

$$\mathrm{Tr}_{A_n}[S_{[n]}] = S_{[n-1]} \otimes I_{B_n}, \tag{46}$$

$$\vdots \tag{47}$$

$$\mathrm{Tr}[S_{[1]}] = 1, \tag{48}$$

so that

$$2^{-D_{\min}^{\varepsilon}(\mathcal{N}^{(n)}\|\mathcal{M}^{(n)})}$$

$$= \inf_{S, S_{[n]}, \dots, S_{[1]} \geqslant 0} \left\{ \begin{array}{r} \mathrm{Tr}\left[S^T \Gamma^{\mathcal{M}^{(n)}}\right]: \\ \mathrm{Tr}\left[S^T \Gamma^{\mathcal{N}^{(n)}}\right] \geqslant 1 - \varepsilon, \\ S \leqslant S_{[n]} \otimes I_{B_n}, \\ \mathrm{Tr}_{A_n}\left[S_{[n]}\right] = S_{[n-1]} \otimes I_{B_n}, \\ \vdots \\ \mathrm{Tr}\left[S_{[1]}\right] = 1 \end{array} \right\}. \quad (49)$$

Finally, the full transpose of S corresponds to a legitimate sub co-strategy, and since we are optimizing over all of them, we can remove the transpose in the optimization to arrive at (39).

The dual program is then given by (40), which can be verified by means of the Lagrange multiplier method. The details of this calculation are provided in the Appendix B 3.

Proposition 4. For strategies $\mathcal{N}^{(n)}$ and $\mathcal{M}^{(n)}$ taking systems A^n to B^n, the distinguishability cost is computable via the following semidefinite program:

$$2^{D_{\max}^{\varepsilon}(\mathcal{N}^{(n)}\|\mathcal{M}^{(n)})}$$

$$= \inf_{\substack{\lambda, Y_n, N \geqslant 0, \\ N_{[n]}, \dots, N_{[1]} \geqslant 0 \\ Y_1, \dots, Y_{n-1} \in \mathrm{Herm}}} \left\{ \begin{array}{r} \lambda: \\ N \leqslant \lambda \Gamma^{\mathcal{M}^{(n)}} \\ Y_n \geqslant \Gamma^{\mathcal{N}^{(n)}} - N, \\ Y_{n-1} \otimes I_{A_n} \geqslant \mathrm{Tr}_{B_n}[Y_n], \\ Y_{n-2} \otimes I_{A_{n-1}} \geqslant \mathrm{Tr}_{B_{n-1}}[Y_{n-1}], \\ \vdots \\ Y_1 \otimes I_{A_2} \geqslant \mathrm{Tr}_{B_2}[Y_2], \\ \varepsilon I_{A_1} \geqslant \mathrm{Tr}_{B_1}[Y_1], \\ \mathrm{Tr}_{B_n}[N] = N_{[n-1]} \otimes I_{A_n}, \\ \mathrm{Tr}_{B_{n-1}}\left[N_{[n-1]}\right] = N_{[n-2]} \otimes I_{A_{n-1}}, \\ \vdots \\ \mathrm{Tr}_{B_2}\left[N_{[2]}\right] = N_{[1]} \otimes I_{A_2}, \\ \mathrm{Tr}_{B_1}\left[N_{[1]}\right] = I_{A_1} \end{array} \right\}, \quad (50)$$

with dual

$$\sup_{W_1, W_2, \dots, \ W_{n+2} \geqslant 0, \ X_1, \dots, X_n \in \mathrm{Herm}}$$

$$\left\{ \begin{array}{r} \mathrm{Tr}[\Gamma^{\mathcal{N}^{(n)}} W_{n+1} - \varepsilon W_1 + X_1]: \\ \mathrm{Tr}[W_{n+2} \Gamma^{\mathcal{M}^{(n)}}] \leqslant 1, \\ W_{n+2} \geqslant W_{n+1} + X_n \otimes I_{B_n}, \\ W_n \otimes I_{B_n} \geqslant W_{n+1}, \\ W_{n-1} \otimes I_{B_{n-1}} \geqslant \mathrm{Tr}_{A_n}[W_n], \\ W_{n-2} \otimes I_{B_{n-2}} \geqslant \mathrm{Tr}_{A_{n-1}}[W_{n-1}], \\ \vdots \\ W_1 \otimes I_{B_1} \geqslant \mathrm{Tr}_{A_2}[W_2], \\ \mathrm{Tr}_{A_n}[X_n] \geqslant X_{n-1} \otimes I_{B_{n-1}}, \\ \mathrm{Tr}_{A_{n-1}}[X_{n-1}] \geqslant X_{n-2} \otimes I_{B_{n-2}}, \\ \vdots \\ \mathrm{Tr}_{A_3}[X_3] \geqslant X_2 \otimes I_{B_2}, \\ \mathrm{Tr}_{A_2}[X_2] \geqslant X_1 \otimes I_{B_1} \end{array} \right\}. \quad (51)$$

Proof. Firstly, we have

$$D_{\max}^{\varepsilon}(\mathcal{N}^{(n)}\|\mathcal{M}^{(n)}) = \inf_{\widetilde{\mathcal{N}}^{(n)} \approx_{\varepsilon} \mathcal{N}^{(n)}} D_{\max}(\widetilde{\mathcal{N}}^{(n)}\|\mathcal{M}^{(n)}) \quad (52)$$

and the dual of the normalized strategy distance from (13)

$$\frac{1}{2}\|\mathcal{N}^{(n)} - \widetilde{\mathcal{N}}^{(n)}\|_{\diamond n}$$

$$= \inf_{\substack{\mu \in \mathbb{R}, Y_n \geqslant 0, \\ Y_1 \dots Y_{n-1} \in \mathrm{Herm}}} \left\{ \begin{array}{r} \mu: \\ Y_n \geqslant \Gamma^{\mathcal{N}^{(n)}} - \Gamma^{\widetilde{\mathcal{N}}^{(n)}}, \\ Y_{n-1} \otimes I_{A_n} \geqslant \mathrm{Tr}_{B_n}[Y_n], \\ Y_{n-2} \otimes I_{A_{n-1}} \geqslant \mathrm{Tr}_{B_{n-1}}[Y_{n-1}], \\ \vdots \\ Y_1 \otimes I_{A_2} \geqslant \mathrm{Tr}_{B_2}[Y_2], \\ \mu I_{A_1} \geqslant \mathrm{Tr}_{B_1}[Y_1] \end{array} \right\}. \quad (53)$$

For $\widetilde{\mathcal{N}}^{(n)}$ the optimizer in (52) and exploiting (36), we have

$$2^{D_{\max}^{\varepsilon}(\mathcal{N}^{(n)}\|\mathcal{M}^{(n)})} = \inf_{\lambda \geqslant 0}\{\lambda: \Gamma^{\widetilde{\mathcal{N}}^{(n)}} \leqslant \lambda \Gamma^{\mathcal{M}^{(n)}}\}. \quad (54)$$

Now we combine these while also adding constraints that ensure that $\widetilde{\mathcal{N}}^{(n)}$ is a valid quantum strategy. Therefore we use the constraints in (5) and incorporate them into the optimization. Thus we get

$$2^{D_{\max}^{\varepsilon}(\mathcal{N}^{(n)}\|\mathcal{M}^{(n)})}$$

$$= \inf_{\substack{\lambda, N, Y_n \geqslant 0, \\ Y_1, \dots, Y_{n-1} \in \mathrm{Herm} \\ N_{[1]}, \dots, N_{[n-1]} \geqslant 0}} \left\{ \begin{array}{r} \lambda: \\ N \leqslant \lambda \Gamma^{\mathcal{M}^{(n)}}, \\ Y_n \geqslant \Gamma^{\mathcal{N}^{(n)}} - N, \\ Y_{n-1} \otimes I_{A_n} \geqslant \mathrm{Tr}_{B_n}[Y_n], \\ Y_{n-2} \otimes I_{A_{n-1}} \geqslant \mathrm{Tr}_{B_{n-1}}[Y_{n-1}], \\ \vdots \\ Y_1 \otimes I_{A_2} \geqslant \mathrm{Tr}_{B_2}[Y_2], \\ \varepsilon I_{A_1} \geqslant \mathrm{Tr}_{B_1}[Y_1], \\ \mathrm{Tr}_{B_n}[N] = N_{[n-1]} \otimes I_{A_n}, \\ \mathrm{Tr}_{B_{n-1}}\left[N_{[n-1]}\right] = N_{[n-2]} \otimes I_{A_{n-1}}, \\ \vdots \\ \mathrm{Tr}_{B_2}\left[N_{[2]}\right] = N_{[1]} \otimes I_{A_2}, \\ \mathrm{Tr}_{B_1}\left[N_{[1]}\right] = I_{A_1} \end{array} \right\}.$$

The dual is given by (51), which can be verified by means of the Lagrange multiplier method accompanying this paper and is worked out in the Appendix B 4. ∎

These SDPs are extensions of those presented in Ref. [14, Appendix C-3], with the difference being that those above incorporate causality constraints for quantum strategies. They enable us to efficiently compute these quantities for various scenarios of interest. We do so in the following section, where we investigate whether adaptive strategies provide an advantage over parallel ones with respect to the quantum strategy distance, the approximate distillable distinguishability, and the approximate distinguishability cost.

IV. ADAPTIVE VS NONADAPTIVE IN DISCRIMINATION AND DISTILLATION

Strategies that distinguish between two channels $\mathcal{N}_{A \to B}$ and $\mathcal{M}_{A \to B}$ using each channel n times are adaptive in general. Parallel strategies are a special case of adaptive strategies that are of practical interest. Parallel strategies involve a distinguisher inputting a possibly entangled state simultaneously to
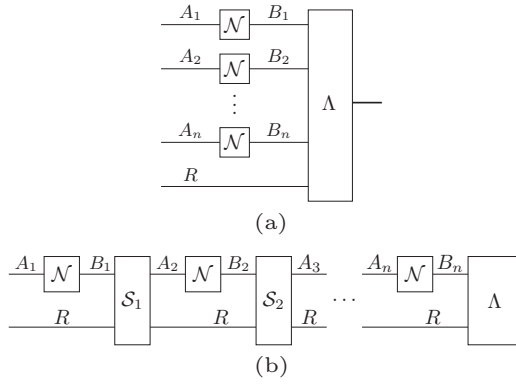
(a)

(b)

FIG. 2. Processing n uses of a channel $\mathcal{N}_{A \to B}$ in (a) a parallel manner and (b) an adaptive manner. In a parallel strategy, the input state to the n copies of the channel can be entangled, but the channels are all applied simultaneously. In an adaptive strategy, the channels are applied sequentially, with interleaving channels $\mathcal{S}_1$ through $\mathcal{S}_{n-1}$ providing adaptive feedback. In a channel discrimination protocol, channels $\mathcal{N}_{A \to B}$ and $\mathcal{M}_{A \to B}$ are processed using a common apparatus and the final measurement results are used to decide the correct channel.

n instances of the unknown channel. Adaptive strategies, on the other hand, involve n uses of the unknown channel that happen sequentially. Between uses of the unknown channel, the distinguisher can perform a quantum channel so as to boost the chances of success. These two scenarios are described in Fig. 2.

A parallel strategy is a special case of an adaptive strategy [19]. Adaptive strategies are therefore no less powerful than parallel ones. It is known that in the asymptotic regime, adaptive strategies confer no advantage over nonadaptive ones in asymmetric channel discrimination [21–23]. However, in practical situations of interest with a finite number of uses of the unknown channel and specific distinguishability tasks, it is possible that adaptive strategies offer an advantage.

The formulation of quantum strategies offers a powerful framework in which to analyze this problem. Consider a strategy $\mathcal{N}^{(n)}$ such as the one in Fig. 1 that consists of n uses of the channel $\mathcal{N}_{A \to B}$. This strategy can be made to interact with a general n-turn co-strategy $\mathcal{S}^{(n)}$, which encapsulates all possible adaptive operations. To study parallel strategies, $\mathcal{N}^{(n)}$ can also be made to interact with a constrained, parallel n-turn co-strategy. These two cases are described in Fig. 2.

In this work we study the gap between adaptive and parallel strategies for distinguishability tasks involving two different generalized amplitude damping channels [33]. The GADC is a qubit-to-qubit channel that is characterized by a damping parameter $\gamma \in [0, 1]$ and a noise parameter $N \in [0, 1]$. It models the dynamics of a qubit system that is in contact with a thermal bath. It is used to describe some of the noise in superconducting-circuit-based quantum computers [34]. We consider two strategies $\mathcal{N}^{(n)}$ and $\mathcal{M}^{(n)}$ that each consist of n uses of a particular GADC. The Choi operator of a GADC $\mathcal{A}_{\gamma,N}$ with damping parameter γ and noise parameter N is
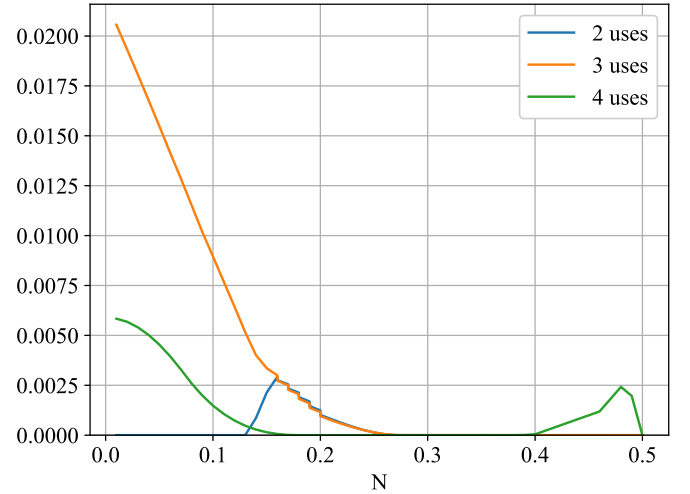

FIG. 3. Consider two GADCs with $\gamma = 0.2$ and 0.3, respectively. We plot the *difference* between $\frac{1}{2}\|\mathcal{N}^{(n)} - \mathcal{M}^{(n)}\|_{\diamond n}$ and $\frac{1}{2}\|\mathcal{N}^{\otimes n} - \mathcal{M}^{\otimes n}\|_{\diamond}$, where the strategies $\mathcal{N}^{(n)}$ and $\mathcal{M}^{(n)}$ each consist of n instances of the same channel. While varying the common parameter N, and allowing for different number of channel uses, we see that adaptive strategies offer an advantage in discrimination over parallel ones.

given by

$$\Gamma_{RB}^{\mathcal{A}_{\gamma,N}} := \begin{bmatrix} 1 - \gamma N & 0 & 0 & \sqrt{1-\gamma} \\ 0 & \gamma N & 0 & 0 \\ 0 & 0 & \gamma(1-N) & 0 \\ \sqrt{1-\gamma} & 0 & 0 & 1 - \gamma(1-N) \end{bmatrix}. \tag{55}$$

In Fig. 3 we plot the *difference* between the strategy distance $\frac{1}{2}\|\mathcal{N}^{(n)} - \mathcal{M}^{(n)}\|_{\diamond n}$ and the diamond distance $\frac{1}{2}\|\mathcal{N}^{\otimes n} - \mathcal{M}^{\otimes n}\|_{\diamond}$. We consider two GADCs with damping parameter $\gamma = 0.2$ and 0.3, respectively, while varying their common noise parameter N. We note here that the strategy distance $\frac{1}{2}\|\mathcal{N}^{(n)} - \mathcal{M}^{(n)}\|_{\diamond n}$ involves an optimization over all costrategies, whereas the optimization involved in the diamond distance $\frac{1}{2}\|\mathcal{N}^{\otimes n} - \mathcal{M}^{\otimes n}\|_{\diamond}$ is restricted to parallel costrategies. This enables us to investigate if adaptive strategies offer an advantage over parallel ones in channel discrimination, and we indeed see in Fig. 3 that there is a nonzero gap between the strategy distance and the diamond distance.

Further, for two GADCs we study the gap between adaptive and parallel costrategies for the distillable distinguishability, as well as the distinguishability cost. We consider that both channels have $\gamma = 0.2$. The channels differ in their noise parameter N, with the first channel having $N = 0.2$ and the second channel having $N = 0.3$. We use the SDP formulations of the smooth strategy min-relative entropy in (39) and the smooth strategy max-relative entropy in (50) to perform this calculation. For the distillable distinguishability, the results of the calculation are in Fig. 4, where we see that there is a gap in the distillable distinguishability between adaptive and parallel strategies. The results of the distinguishability cost are in Fig. 5, where again we see a nonzero gap between adaptive and parallel strategies. The code used to produce
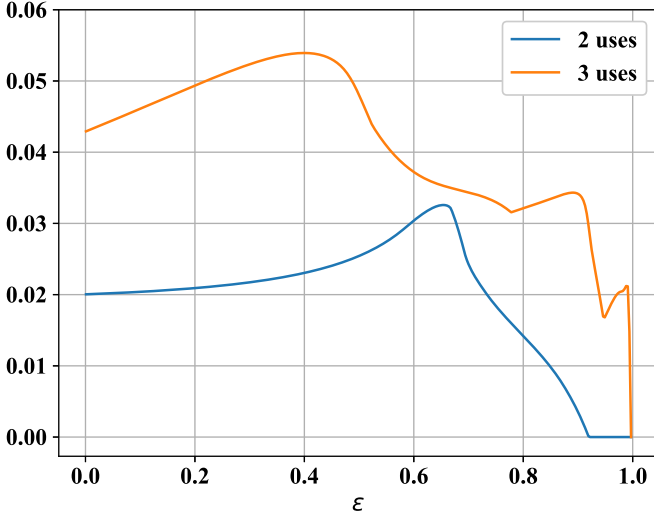
FIG. 4. Consider two GADCs with $N = 0.2$ and 0.3, respectively. They both have $\gamma = 0.2$. We plot the *difference* between the distillable distinguishabilities, which is given by $D_{\min}^{\varepsilon}(\mathcal{N}^{(n)} \| \mathcal{M}^{(n)})$ for the most general case and by $D_{\min}^{\varepsilon}(\mathcal{N}^{\otimes n} \| \mathcal{M}^{\otimes n})$ for the case when a distinguisher is limited to a parallel strategy.

these numerical results is available with the arXiv version of the paper. It is not yet clear to us how to explain this behavior qualitatively, and so we leave it for future work to do so.

V. CONCLUSION

In summary, in this paper we reviewed and further developed the resource theory of asymmetric distinguishability for quantum strategies, which is a high-level and flexible framework with which to study quantum interactions. We provided semidefinite programs to calculate the distillable
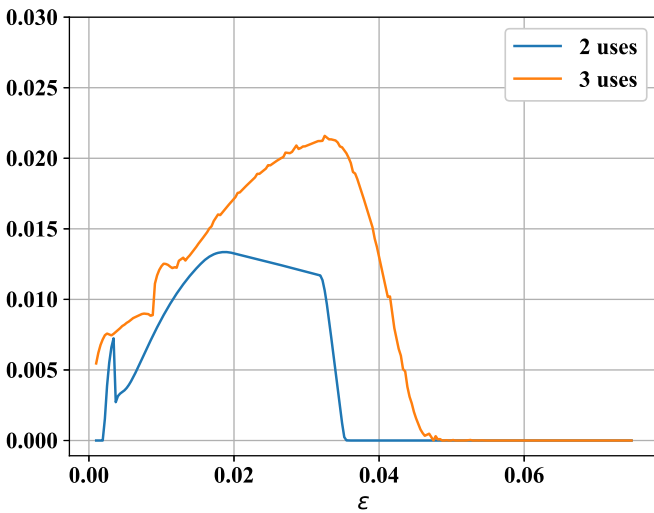

FIG. 5. Consider two GADCs with $N = 0.2$ and 0.3, respectively. They both have $\gamma = 0.2$. We plot the *difference* between the approximate distinguishability cost, which is given by $D_{\max}^{\varepsilon}(\mathcal{N}^{(n)} \| \mathcal{M}^{(n)})$ for the most general case and by $D_{\max}^{\varepsilon}(\mathcal{N}^{\otimes n} \| \mathcal{M}^{\otimes n})$ for the case when an adversary is limited to a parallel strategy.

distinguishability and the distinguishability cost of quantum strategy boxes, which we used to compare the power of adaptive strategies to parallel ones. It is known that for channel discrimination and distillable distinguishability, parallel strategies are equally powerful as adaptive strategies in the asymptotic limit; however, an example we considered shows that adaptive strategies provide an advantage in general when one considers a finite number of channel uses.

Note added. There has been further work conducted on this topic recently, with Refs. [35,36] studying the difference between adaptive and parallel strategies in channel distinguishability tasks, and Refs. [37,38] studying the task of discriminating between amplitude damping channels.

APPENDIX A: ONE-SHOT DISTILLATION AND DILUTION OF STRATEGY BOXES

In the following we provide the proof of Theorem 2, which is claimed in [14]. For completeness, we restate the theorem below:

Theorem 5. In Ref. [14] the approximate one-shot distillable distinguishability of the strategy box $(\mathcal{N}^{(n)}, \mathcal{M}^{(n)})$ is equal to the smooth strategy min-relative entropy:

$$D_d^{\varepsilon}(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) = D_{\min}^{\varepsilon}(\mathcal{N}^{(n)} \| \mathcal{M}^{(n)}), \quad (A1)$$

and the approximate one-shot distinguishability cost is equal to the smooth strategy max-relative entropy:

$$D_c^{\varepsilon}(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) = D_{\max}^{\varepsilon}(\mathcal{N}^{(n)} \| \mathcal{M}^{(n)}). \quad (A2)$$

1. One-shot exact distillable distinguishability is strategy min-relative entropy

We first prove the inequality

$$D_d^0(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) \geqslant D_{\min}(\mathcal{N}^{(n)} \| \mathcal{M}^{(n)}). \quad (A3)$$

Let Θ be an arbitrary n-turn co-strategy that interacts with strategies $\mathcal{N}^{(n)}$ or $\mathcal{M}^{(n)}$ to yield a state on $R_n B_n$. Consider the projector $0 \leqslant \Lambda_{R_n B_n} \leqslant I_{R_n B_n}$ onto the support of $\Theta \circ \mathcal{N}^{(n)}$. Consider a postprocessing of the output state $\omega_{R_n B_n}$ as follows:

$$\omega_{R_n B_n} \rightarrow \mathrm{Tr}\left[\Lambda_{R_n B_n} \omega_{R_n B_n}\right] |0\rangle\langle 0|_X$$
$$+ \mathrm{Tr}\left[(I_{R_n B_n} - \Lambda_{R_n B_n}) \omega_{R_n B_n}\right] |1\rangle\langle 1|_X. \quad (A4)$$

If the unknown strategy is $\mathcal{N}^{(n)}$, then the interaction with Θ followed by the above postprocessing yields $|0\rangle\langle 0|_X$. If the unknown strategy is $\mathcal{M}^{(n)}$, then the final state is π_M with

$$M = \frac{1}{\mathrm{Tr}\left[\Lambda_{R_n B_n}(\Theta \circ \mathcal{M}^{(n)})\right]}, \quad (A5)$$

or equivalently,

$$\log_2 M = D_{\min}(\Theta \circ \mathcal{N}^{(n)} \| \Theta \circ \mathcal{M}^{(n)}). \quad (A6)$$

Taking a supremum over all interacting costrategies Θ, we get

$$D_d^0(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) \geqslant \sup_\Theta D_{\min}(\Theta \circ \mathcal{N}^{(n)} \| \Theta \circ \mathcal{M}^{(n)}) \quad \text{(A7)}$$

$$= D_{\min}(\mathcal{N}^{(n)} \| \mathcal{M}^{(n)}). \quad \text{(A8)}$$

Next we prove the opposite inequality,

$$D_d^0(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) \leqslant D_{\min}(\mathcal{N}^{(n)} \| \mathcal{M}^{(n)}), \quad \text{(A9)}$$

which is a consequence of the data-processing inequality for the $D_{\min}$ strategy divergence [14]. Consider an arbitrary n-turn co-strategy Θ that interacts with $\mathcal{N}^{(n)}$ to give $|0\rangle\langle 0|$ and with $\mathcal{M}^{(n)}$ to give π_M. Then we can write

$$D_{\min}(\mathcal{N}^{(n)} \| \mathcal{M}^{(n)}) \geqslant D_{\min}(\Theta \circ \mathcal{N}^{(n)} \| \Theta \circ \mathcal{M}^{(n)}) \quad \text{(A10)}$$

$$= D_{\min}(|0\rangle\langle 0| \| \pi_M) \quad \text{(A11)}$$

$$= \log_2 M, \quad \text{(A12)}$$

which yields

$$D_{\min}(\mathcal{N}^{(n)} \| \mathcal{M}^{(n)}) \geqslant D_d^0(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}). \quad \text{(A13)}$$

Putting (A8) and (A13) together, we get

$$D_d^0(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) = D_{\min}(\mathcal{N}^{(n)} \| \mathcal{M}^{(n)}). \quad \text{(A14)}$$

2. One-shot approximate distillable distinguishability is smooth strategy min-relative entropy

Here our aim is to prove

$$D_d^\varepsilon(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) = D_{\min}^\varepsilon(\mathcal{N}^{(n)} \| \mathcal{M}^{(n)}). \quad \text{(A15)}$$

First we prove the inequality

$$D_d^\varepsilon(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) \geqslant D_{\min}^\varepsilon(\mathcal{N}^{(n)} \| \mathcal{M}^{(n)}). \quad \text{(A16)}$$

Let Θ be an arbitrary interacting n-turn co-strategy and $\Lambda_{R_n B_n}$ a corresponding measurement operator satisfying $0 \leqslant \Lambda_{R_n B_n} \leqslant I_{R_n B_n}$ and

$$\text{Tr}[\Lambda_{R_n B_n}(\Theta \circ \mathcal{N}^{(n)})] \geqslant 1 - \varepsilon. \quad \text{(A17)}$$

Consider, as in the exact case, a postprocessing of the final state $\omega_{R_n B_n}$ by the measurement channel $\mathcal{L}_{R_n R_n \to X}$:

$$\mathcal{L}_{R_n R_n \to X}(\omega_{RB^n}) := \text{Tr}[\Lambda_{R_n B_n} \omega_{R_n B_n}] |0\rangle\langle 0|_X$$
$$+ \text{Tr}[(I_{R_n B_n} - \Lambda_{R_n B_n}) \omega_{R_n B_n}] |1\rangle\langle 1|_X. \quad \text{(A18)}$$

Using (A17), we can conclude that $\mathcal{L} \circ \Theta \circ \mathcal{N}^{(n)} \approx_\varepsilon |0\rangle\langle 0|$. Further, for

$$M = \frac{1}{\text{Tr}[\Lambda_{R_n B_n}(\Theta \circ \mathcal{M}^{(n)})]}, \quad \text{(A19)}$$

we have $\mathcal{L} \circ \Theta \circ \mathcal{M}^{(n)} = \pi_M$. Taking a supremum over all interacting costrategies Θ and measurement channels $\mathcal{L}_{R_n R_n \to X}$, we get

$$D_d^\varepsilon(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) \geqslant \sup_\Theta D_{\min}^\varepsilon(\Theta \circ \mathcal{N}^{(n)} \| \Theta \circ \mathcal{M}^{(n)})$$

$$= D_{\min}^\varepsilon(\mathcal{N}^{(n)} \| \mathcal{M}^{(n)}). \quad \text{(A20)}$$

Next we use data processing to prove the reverse inequality

$$D_d^\varepsilon(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) \leqslant D_{\min}^\varepsilon(\mathcal{N}^{(n)} \| \mathcal{M}^{(n)}). \quad \text{(A21)}$$

Consider an n-turn co-strategy Θ and measurement channel $\mathcal{L}_{R_n R_n \to X}$ such that $\frac{1}{2}\|\mathcal{L} \circ \Theta \circ \mathcal{N}^{(n)} - |0\rangle\langle 0|\|_1 \leqslant \varepsilon$. By a direct calculation with trace distance, we find that

$$\varepsilon \geqslant \frac{1}{2}\|\mathcal{L} \circ \Theta \circ \mathcal{N}^{(n)} - |0\rangle\langle 0|\|_1 \quad \text{(A22)}$$

$$= 1 - \text{Tr}[\Lambda(\Theta \circ \mathcal{N}^{(n)})]. \quad \text{(A23)}$$

We conclude that $\text{Tr}[\Lambda(\Theta \circ \mathcal{N}^{(n)})] \geqslant 1 - \varepsilon$. In the definition of $D_{\min}^\varepsilon(\Theta \circ \mathcal{N}^{(n)} \| \pi_M)$, we can take the final measurement operator to be Λ_{RB}. This leaves us with $\text{Tr}[\Lambda_{R_n B_n}(\Theta \circ \mathcal{N}^{(n)})] \geqslant 1 - \varepsilon$ and $\text{Tr}[\Lambda_{R_n B_n}(\Theta \circ \mathcal{M}^{(n)})] = 1/M$. Since the definition of $D_{\min}^\varepsilon$ for strategies involves an optimization over costrategies and measurement operators, we conclude that

$$D_{\min}^\varepsilon(\mathcal{N}^{(n)} \| \mathcal{M}^{(n)}) \geqslant D_{\min}^\varepsilon(\mathcal{L} \circ \Theta \circ \mathcal{N}^{(n)} \| \pi_M) \quad \text{(A24)}$$

$$\geqslant \log_2 M, \quad \text{(A25)}$$

where the last inequality follows from Ref. [13, Appendix F-1]. Since the scheme considered for distillation is arbitrary, we conclude that

$$D_{\min}^\varepsilon(\mathcal{N}^{(n)} \| \mathcal{M}^{(n)}) \geqslant D_d^\varepsilon(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}). \quad \text{(A26)}$$

Combining (A20) and (A26), we obtain the desired result:

$$D_d^\varepsilon(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) = D_{\min}^\varepsilon(\mathcal{N}^{(n)} \| \mathcal{M}^{(n)}). \quad \text{(A27)}$$

3. One-shot exact distinguishability cost is strategy max-relative entropy

First we aim to prove the inequality

$$D_c^0(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) \leqslant D_{\max}(\mathcal{N}^{(n)} \| \mathcal{M}^{(n)}). \quad \text{(A28)}$$

To do so, we first let λ be such that

$$\mathcal{N}^{(n)} \leqslant 2^\lambda \mathcal{M}^{(n)}. \quad \text{(A29)}$$

This means that

$$\mathcal{N}'^{(n)} := \frac{2^\lambda \mathcal{M}^{(n)} - \mathcal{N}^{(n)}}{2^\lambda - 1} \quad \text{(A30)}$$

is a quantum strategy. Further, if the Choi operators of $\mathcal{N}^{(n)}$ and $\mathcal{M}^{(n)}$ are $\Gamma^{\mathcal{N}^{(n)}}$ and $\Gamma^{\mathcal{M}^{(n)}}$, respectively, then

$$\frac{2^\lambda \Gamma^{\mathcal{M}^{(n)}} - \Gamma^{\mathcal{N}^{(n)}}}{2^\lambda - 1} \quad \text{(A31)}$$

is the Choi operator of $\mathcal{N}'^{(n)}$ (by linearity).

Consider an arbitrary n-turn co-strategy that is made to act as follows, beginning with system X. It acts as follows:

$$\sigma_X \to (\Theta \circ \mathcal{N}^{(n)}) \langle 0| \sigma_X |0\rangle + (\Theta \circ \mathcal{N}'^{(n)}) \langle 1| \sigma_X |1\rangle. \quad \text{(A32)}$$

In the case that $\sigma_X = |0\rangle\langle 0|_X$, then the output is $\Theta \circ \mathcal{N}^{(n)}$. If the input is π_M where $M = 2^\lambda$, then the output is $\Theta \circ \mathcal{M}^{(n)}$.

For this particular choice of transformation, we obtain a distinguishability cost of λ, so if one optimizes over all protocols, one obtains $D_c^0(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) \leqslant \lambda$. Now if we optimize over all λ such that (A29) holds, we obtain

$$D_c^0(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) \leqslant D_{\max}(\mathcal{N}^{(n)} \| \mathcal{M}^{(n)}). \quad \text{(A33)}$$

The opposite inequality follows from the data processing inequality for the strategy max-relative entropy [14]. Let Θ be a strategy satisfying

$$\Theta(|0\rangle\langle 0|) = \mathcal{N}^{(n)}, \text{ and} \tag{A34}$$

$$\Theta(\pi_M) = \mathcal{M}^{(n)}, \tag{A35}$$

with $\log_2 M = D_c^0(\mathcal{N}^{(n)}, \mathcal{M}^{(n)})$. Then consider the following chain of reasoning:

$$\log_2 M = D_c^0(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) \tag{A36}$$

$$= D_{\max}(|0\rangle\langle 0\|\pi_M) \tag{A37}$$

$$\geqslant D_{\max}(\Theta(|0\rangle\langle 0|)\|\Theta(\pi_M)) \tag{A38}$$

$$= D_{\max}(\mathcal{N}^{(n)}\|\mathcal{M}^{(n)}). \tag{A39}$$

This lets us conclude that

$$D_c^0(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) \geqslant D_{\max}(\mathcal{N}^{(n)}\|\mathcal{M}^{(n)}). \tag{A40}$$

Putting together (A33) and (A40), we obtain the desired result, which is

$$D_c^0(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) = D_{\max}(\mathcal{N}^{(n)}\|\mathcal{M}^{(n)}). \tag{A41}$$

4. One-shot approximate distinguishability cost is smooth strategy max-relative entropy

Here we aim to prove that

$$D_c^\varepsilon(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) = D_{\max}^\varepsilon(\mathcal{N}^{(n)}\|\mathcal{M}^{(n)}). \tag{A42}$$

First, we prove the inequality

$$D_c^\varepsilon(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) \leqslant D_{\max}^\varepsilon(\mathcal{N}^{(n)}\|\mathcal{M}^{(n)}). \tag{A43}$$

To do so, we consider a quantum strategy $\mathcal{N}'^{(n)} \approx_\varepsilon \mathcal{N}^{(n)}$ (which means that $\frac{1}{2}\|\mathcal{N}'^{(n)} - \mathcal{N}^{(n)}\| \leqslant \varepsilon$). We use the construction for the exact distinguishability cost, but instead for $\mathcal{N}'^{(n)}$, and therefore obtain

$$D_c^\varepsilon(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) \leqslant D_{\max}^\varepsilon(\mathcal{N}'^{(n)}\|\mathcal{M}^{(n)}). \tag{A44}$$

By optimizing the above over all $\mathcal{N}'^{(n)}$ satisfying $\mathcal{N}'^{(n)} \approx_\varepsilon \mathcal{N}^{(n)}$, we obtain

$$D_c^\varepsilon(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) \leqslant D_{\max}^\varepsilon(\mathcal{N}^{(n)}\|\mathcal{M}^{(n)}). \tag{A45}$$

To prove the reverse inequality

$$D_c^\varepsilon(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) \geqslant D_{\max}^\varepsilon(\mathcal{N}^{(n)}\|\mathcal{M}^{(n)}), \tag{A46}$$

we again use data-processing arguments. Consider first a strategy Θ such that

$$\Theta(|0\rangle\langle 0|) \approx_\varepsilon \mathcal{N}^{(n)}, \text{ and} \tag{A47}$$

$$\Theta(\pi_M) = \mathcal{M}^{(n)}, \tag{A48}$$

with $\log_2 M = D_c^0(\mathcal{N}^{(n)}, \mathcal{M}^{(n)})$. Now consider the following:

$$D_c^\varepsilon(\mathcal{N}^{(n)}, \mathcal{M}^{(n)}) = \log_2 M \tag{A49}$$

$$= D_{\max}(|0\rangle\langle 0\|\pi_M) \tag{A50}$$

$$\geqslant D_{\max}(\Theta(|0\rangle\langle 0|)\|\Theta(\pi_M)) \tag{A51}$$

$$= D_{\max}(\Theta(|0\rangle\langle 0|)\|\mathcal{M}^{(n)}) \tag{A52}$$

$$\geqslant D_{\max}^\varepsilon(\mathcal{N}^{(n)}\|\mathcal{M}^{(n)}). \tag{A53}$$

Putting together (A45) and (A53), we get the desired result.

APPENDIX B: DERIVATION OF SEMIDEFINITE PROGRAMMING DUALS

Here we provide full details of how to arrive at the duals for the semidefinite programs for the normalized strategy distance, the smooth strategy min-relative entropy, and the smooth strategy max-relative entropy.

1. Background

Suppose that a semidefinite program is given in primal form as follows:

$$\sup_{X \geqslant 0} \{\text{Tr}[AX] : \Phi_1(X) = B_1, \Phi_2(X) \leqslant B_2\}. \tag{B1}$$

Then its dual is given by

$$\inf_{Y_1 \in \text{Herm}, Y_2 \geqslant 0} \left\{ \text{Tr}[B_1 Y_1] + \text{Tr}[B_2 Y_2] : \Phi_1^\dagger(Y_1) + \Phi_2^\dagger(Y_2) \geqslant A \right\}. \tag{B2}$$

We use (B1) and (B2) in the forthcoming sections to derive the various duals presented in our paper.

Alternatively, the following is useful as well. If the primal can be written as

$$\inf_{X \geqslant 0} \{\text{Tr}[(-A)X] : \Phi_1(X) = B_1, \Phi_2(X) \leqslant B_2\}, \tag{B3}$$

then its dual is given by

$$\sup_{Y_1 \in \text{Herm}, Y_2 \geqslant 0} \left\{ -\text{Tr}[B_1 Y_1] - \text{Tr}[B_2 Y_2] : \Phi_1^\dagger(Y_1) + \Phi_2^\dagger(Y_2) \geqslant A \right\}, \tag{B4}$$

which comes about by applying a minus sign to (B1) and carrying it through.

2. Normalized strategy distance

First we repeat the primal for the normalized strategy distance given in (12) in the main text:

$$\sup_{S, S_{[1]}, \ldots, S_{[n]} \geqslant 0} \text{Tr}[S(\Gamma^{\mathcal{N}^{(n)}} - \Gamma^{\mathcal{M}^{(n)}})] \tag{B5}$$

subject to

$$S \leqslant S_{[n]} \otimes I_{B_n}, \tag{B6}$$

$$\text{Tr}_{A_n}\left[S_{[n]}\right] = S_{[n-1]} \otimes I_{B_{n-1}}, \tag{B7}$$

$$\text{Tr}_{A_{n-1}}\left[S_{[n-1]}\right] = S_{[n-2]} \otimes I_{B_{n-2}}, \tag{B8}$$

$$\vdots$$

$$\text{Tr}_{A_2}\left[S_{[2]}\right] = S_{[1]} \otimes I_{B_1}, \tag{B9}$$

$$\text{Tr}\left[S_{[1]}\right] = 1. \tag{B10}$$

Now mapping to (B1), we find that

$$X = \text{diag}(S, S_{[n]}, S_{[n-1]}, \ldots, S_{[2]}, S_{[1]}), \tag{B11}$$

$$A = \text{diag}(\Gamma^{\mathcal{N}^{(n)}} - \Gamma^{\mathcal{M}^{(n)}}, 0, 0, \ldots, 0, 0), \tag{B12}$$

$$\Phi_1(X) = \text{diag}(\text{Tr}_{A_n}[S_{[n]}] - S_{[n-1]} \otimes I_{B_{n-1}},$$

$$\text{Tr}_{A_{n-1}}[S_{[n-1]}] - S_{[n-2]} \otimes I_{B_{n-2}}, \dots,$$

$$\text{Tr}_{A_2}[S_{[2]}] - S_{[1]} \otimes I_{B_1}, \text{Tr}[S_{[1]}]), \tag{B13}$$

$$B_1 = \text{diag}(0, 0, \dots, 0, 1), \tag{B14}$$

$$\Phi_2(X) = S - S_{[n]} \otimes I_{B_n}, \tag{B15}$$

$$B_2 = 0. \tag{B16}$$

We should now figure out the adjoints of Φ_1 and Φ_2. Consider that

$$\text{Tr}[Y_i \Phi_i(X)] = \text{Tr}[\Phi_i^\dagger(Y_i) X] \text{ for } i \in \{1, 2\}. \tag{B17}$$

Set

$$Y_1 = \text{diag}(Z_{n-1}, Z_{n-2}, \dots, Z_1, \mu). \tag{B18}$$

Then we find that

$$\text{Tr}[Y_1 \Phi_1(X)] = \text{Tr}[Z_{n-1}(\text{Tr}_{A_n}[S_{[n]}] - S_{[n-1]} \otimes I_{B_{n-1}})] + \text{Tr}[Z_{n-2}(\text{Tr}_{A_{n-1}}[S_{[n-1]}] - S_{[n-2]} \otimes I_{B_{n-2}})] + \cdots$$

$$+ \text{Tr}[Z_1(\text{Tr}_{A_2}[S_{[2]}] - S_{[1]} \otimes I_{B_1})] + \mu \, \text{Tr}[S_{[1]}] \tag{B19}$$

$$= \text{Tr}[(Z_{n-1} \otimes I_{A_n})S_{[n]}] - \text{Tr}[\text{Tr}_{B_{n-1}}[Z_{n-1}]S_{[n-1]}] + \text{Tr}[(Z_{n-2} \otimes I_{A_{n-1}})S_{[n-1]}] - \text{Tr}[\text{Tr}_{B_{n-2}}[Z_{n-2}]S_{[n-2]}] + \cdots$$

$$+ \text{Tr}[(Z_1 \otimes I_{A_2})S_{[2]}] - \text{Tr}[\text{Tr}_{B_1}[Z_1]S_{[1]}] + \text{Tr}[\mu I_{A_1}S_{[1]}] \tag{B20}$$

$$= \text{Tr}[(Z_{n-1} \otimes I_{A_n})S_{[n]}] + \text{Tr}[(Z_{n-2} \otimes I_{A_{n-1}} - \text{Tr}_{B_{n-1}}[Z_{n-1}])S_{[n-1]}] + \text{Tr}[(Z_{n-3} \otimes I_{A_{n-2}} - \text{Tr}_{B_{n-2}}[Z_{n-2}])S_{[n-2]}] + \cdots$$

$$+ \text{Tr}[(Z_1 \otimes I_{A_2} - \text{Tr}_{B_2}[Z_2])S_{[2]}] + \text{Tr}[(\mu I_{A_1} - \text{Tr}_{B_1}[Z_1])S_{[1]}]. \tag{B21}$$

So this implies that

$$\Phi_1^\dagger(Y_1) = \text{diag}(0, Z_{n-1} \otimes I_{A_n}, Z_{n-2} \otimes I_{A_{n-1}} - \text{Tr}_{B_{n-1}}[Z_{n-1}], \dots, Z_1 \otimes I_{A_2} - \text{Tr}_{B_2}[Z_2], \mu I_{A_1} - \text{Tr}_{B_1}[Z_1]). \tag{B22}$$

Now set

$$Y_2 = Z_n, \tag{B23}$$

and we find that

$$\text{Tr}[Y_2 \Phi_2(X)] = \text{Tr}[Z_n(S - S_{[n]} \otimes I_{B_n})] \tag{B24}$$

$$= \text{Tr}[Z_n S] - \text{Tr}[\text{Tr}_{B_n}[Z_n]S_{[n]}], \tag{B25}$$

so that

$$\Phi_2^\dagger(Y_2) = \text{diag}(Z_n, -\text{Tr}_{B_n}[Z_n], 0, \dots, 0, 0). \tag{B26}$$

We finally find that

$$\Phi_1^\dagger(Y_1) + \Phi_2^\dagger(Y_2) = \text{diag}(Z_n, Z_{n-1} \otimes I_{A_n} - \text{Tr}_{B_n}[Z_n], Z_{n-2} \otimes I_{A_{n-1}} - \text{Tr}_{B_{n-1}}[Z_{n-1}], \dots, Z_1 \otimes I_{A_2} - \text{Tr}_{B_2}[Z_2], \mu I_{A_1} - \text{Tr}_{B_1}[Z_1]), \tag{B27}$$

and so $\Phi_1^\dagger(Y_1) + \Phi_2^\dagger(Y_2) \geqslant A$ is equivalent to the following conditions:

$$Z_n \geqslant \Gamma^{\mathcal{N}^{(n)}} - \Gamma^{\mathcal{M}^{(n)}}, \tag{B28}$$

$$Z_{n-1} \otimes I_{A_n} \geqslant \text{Tr}_{B_n}[Z_n], \tag{B29}$$

$$Z_{n-2} \otimes I_{A_{n-1}} \geqslant \text{Tr}_{B_{n-1}}[Z_{n-1}], \tag{B30}$$

$$\vdots$$

$$Z_1 \otimes I_{A_2} \geqslant \text{Tr}_{B_2}[Z_2], \tag{B31}$$

$$\mu I_{A_1} \geqslant \text{Tr}_{B_1}[Z_1]. \tag{B32}$$

So then the dual is given by plugging into (B2):

$$\inf_{\substack{Z_{n-1}, Z_{n-2}, \dots, Z_1 \in \text{Herm}, \\ \mu \in \mathbb{R}, Z_n \geqslant 0}} \mu, \tag{B33}$$

subject to

$$Z_n \geqslant \Gamma^{\mathcal{N}^{(n)}} - \Gamma^{\mathcal{M}^{(n)}}, \tag{B34}$$

$$Z_{n-1} \otimes I_{A_n} \geqslant \text{Tr}_{B_n}[Z_n], \tag{B35}$$

$$Z_{n-2} \otimes I_{A_{n-1}} \geqslant \text{Tr}_{B_{n-1}}[Z_{n-1}], \tag{B36}$$

$$\vdots$$

$$Z_1 \otimes I_{A_2} \geqslant \text{Tr}_{B_2}[Z_2], \tag{B37}$$

$$\mu I_{A_1} \geqslant \text{Tr}_{B_1}[Z_1]. \tag{B38}$$

3. Smooth strategy min-relative entropy

First we repeat the primal for the smooth strategy min-relative entropy given in (39) in the main text:

$$\inf_{S, S_{[1]}, \dots, S_{[n]} \geqslant 0} \text{Tr}[S \, \Gamma^{\mathcal{M}^{(n)}}] \tag{B39}$$

subject to

$$\text{Tr}[S \, \Gamma^{\mathcal{N}^{(n)}}] \geqslant 1 - \varepsilon \tag{B40}$$

$$S \leqslant S_{[n]} \otimes I_{B_n}, \tag{B41}$$

$$\text{Tr}_{A_n}[S_{[n]}] = S_{[n-1]} \otimes I_{B_{n-1}}, \quad \text{(B42)}$$

$$\text{Tr}_{A_{n-1}}\left[S_{[n-1]}\right] = S_{[n-2]} \otimes I_{B_{n-2}}, \quad \text{(B43)}$$

$$\vdots$$

$$\text{Tr}_{A_2}\left[S_{[2]}\right] = S_{[1]} \otimes I_{B_1}, \quad \text{(B44)}$$

$$\text{Tr}\left[S_{[1]}\right] = 1. \quad \text{(B45)}$$

Now mapping to (B3), we find that

$$X = \text{diag}(S, S_{[n]}, S_{[n-1]}, \dots, S_{[2]}, S_{[1]}), \quad \text{(B46)}$$

$$A = \text{diag}(-\Gamma^{\mathcal{M}^{(n)}}, 0, 0, \dots, 0, 0), \quad \text{(B47)}$$

$$\Phi_1(X) = \text{diag}(\text{Tr}_{A_n}\left[S_{[n]}\right] - S_{[n-1]} \otimes I_{B_{n-1}},$$

$$\text{Tr}_{A_{n-1}}\left[S_{[n-1]}\right] - S_{[n-2]} \otimes I_{B_{n-2}}, \dots,$$

$$\text{Tr}_{A_2}\left[S_{[2]}\right] - S_{[1]} \otimes I_{B_1}, \text{Tr}\left[S_{[1]}\right]), \quad \text{(B48)}$$

$$B_1 = \text{diag}(0, 0, \dots, 0, 1), \quad \text{(B49)}$$

$$\Phi_2(X) = \text{diag}(- \text{Tr}[S \, \Gamma^{\mathcal{N}^{(n)}}], S - S_{[n]} \otimes I_{B_n}), \quad \text{(B50)}$$

$$B_2 = \text{diag}(-(1 - \varepsilon), 0). \quad \text{(B51)}$$

The variable Y_1 is the same as in (B18), and the adjoint of Φ_1 is the same as in (B22). Let us set

$$Y_2 = \text{diag}(\mu_1, Z_n). \quad \text{(B52)}$$

Then we find that

$$\text{Tr}[Y_2 \Phi_2(X)] = -\mu_1 \text{Tr}[S \, \Gamma^{\mathcal{N}^{(n)}}] + \text{Tr}[Z_n(S - S_{[n]} \otimes I_{B_n})] \quad \text{(B53)}$$

$$= \text{Tr}[(Z_n - \mu_1 \Gamma^{\mathcal{N}^{(n)}})S] - \text{Tr}[\text{Tr}_{B_n}[Z_n]S_{[n]}], \quad \text{(B54)}$$

so that

$$\Phi_2^\dagger(Y_2) = \text{diag}(Z_n - \mu_1 \Gamma^{\mathcal{N}^{(n)}}, - \text{Tr}_{B_n}[Z_n], 0, \dots, 0, 0). \quad \text{(B55)}$$

So then we find that

$$\Phi_1^\dagger(Y_1) + \Phi_2^\dagger(Y_2) = \text{diag}(Z_n - \mu_1 \Gamma^{\mathcal{N}^{(n)}}, Z_{n-1} \otimes I_{A_n}$$

$$- \text{Tr}_{B_n}[Z_n], Z_{n-2} \otimes I_{A_{n-1}} - \text{Tr}_{B_{n-1}}[Z_{n-1}],$$

$$\dots, Z_1 \otimes I_{A_2} - \text{Tr}_{B_2}[Z_2], \mu I_{A_1}$$

$$- \text{Tr}_{B_1}[Z_1]), \quad \text{(B56)}$$

and so $\Phi_1^\dagger(Y_1) + \Phi_2^\dagger(Y_2) \geqslant A$ is equivalent to the following conditions:

$$Z_n - \mu_1 \Gamma^{\mathcal{N}^{(n)}} \geqslant -\Gamma^{\mathcal{M}^{(n)}}, \quad \text{(B57)}$$

$$Z_{n-1} \otimes I_{A_n} \geqslant \text{Tr}_{B_n}[Z_n], \quad \text{(B58)}$$

$$Z_{n-2} \otimes I_{A_{n-1}} \geqslant \text{Tr}_{B_{n-1}}[Z_{n-1}], \quad \text{(B59)}$$

$$\vdots$$

$$Z_1 \otimes I_{A_2} \geqslant \text{Tr}_{B_2}[Z_2], \quad \text{(B60)}$$

$$\mu I_{A_1} \geqslant \text{Tr}_{B_1}[Z_1]. \quad \text{(B61)}$$

Also, observe that

$$\text{Tr}[B_1 Y_1] + \text{Tr}[B_2 Y_2] = \mu - \mu_1(1 - \varepsilon). \quad \text{(B62)}$$

Thus we conclude after plugging into (B4) that the dual is given by

$$\sup_{\mu_1, Z_n \geqslant 0, \ \mu \in \mathbb{R}, Z_1, \dots, Z_{n-1} \in \text{Herm}} \mu_1(1 - \varepsilon) - \mu \quad \text{(B63)}$$

subject to

$$Z_n \geqslant \mu_1 \Gamma^{\mathcal{N}^{(n)}} - \Gamma^{\mathcal{M}^{(n)}}, \quad \text{(B64)}$$

$$Z_{n-1} \otimes I_{A_n} \geqslant \text{Tr}_{B_n}[Z_n], \quad \text{(B65)}$$

$$Z_{n-2} \otimes I_{A_{n-1}} \geqslant \text{Tr}_{B_{n-1}}[Z_{n-1}], \quad \text{(B66)}$$

$$\vdots$$

$$Z_1 \otimes I_{A_2} \geqslant \text{Tr}_{B_2}[Z_2], \quad \text{(B67)}$$

$$\mu I_{A_1} \geqslant \text{Tr}_{B_1}[Z_1]. \quad \text{(B68)}$$

4. Smooth strategy max-relative entropy

First, we repeat the primal form of the smooth strategy max-relative entropy, given in (50) in the main text:

$$\inf_{\lambda, Y_n, N \geqslant 0, \ N_{[n-1]}, \dots, N_{[1]} \geqslant 0, \ Y_1, \dots, Y_{n-1} \in \text{Herm}} \lambda \quad \text{(B69)}$$

subject to

$$N \leqslant \lambda \Gamma^{\mathcal{M}^{(n)}}, \quad \text{(B70)}$$

$$Y_n \geqslant \Gamma^{\mathcal{N}^{(n)}} - N, \quad \text{(B71)}$$

$$Y_{n-1} \otimes I_{A_n} \geqslant \text{Tr}_{B_n}[Y_n], \quad \text{(B72)}$$

$$Y_{n-2} \otimes I_{A_{n-1}} \geqslant \text{Tr}_{B_{n-1}}[Y_{n-1}], \quad \text{(B73)}$$

$$\vdots$$

$$Y_1 \otimes I_{A_2} \geqslant \text{Tr}_{B_2}[Y_2], \quad \text{(B74)}$$

$$\varepsilon I_{A_1} \geqslant \text{Tr}_{B_1}[Y_1], \quad \text{(B75)}$$

$$\text{Tr}_{B_n}[N] = N_{[n-1]} \otimes I_{A_n}, \quad \text{(B76)}$$

$$\text{Tr}_{B_{n-1}}[N_{[n-1]}] = N_{[n-2]} \otimes I_{A_{n-1}}, \quad \text{(B77)}$$

$$\vdots$$

$$\text{Tr}_{B_2}[N_{[2]}] = N_{[1]} \otimes I_{A_2}, \quad \text{(B78)}$$

$$\text{Tr}_{B_1}[N_{[1]}] = I_{A_1}. \quad \text{(B79)}$$

As a consequence of $Y_n \geqslant 0$ and the constraints above, it follows that $Y_{n-1}, \ldots, Y_1 \geqslant 0$. So the above SDP can be cast in the form of (B3), with

$$X = (\lambda, Y_n, Y_{n-1}, \ldots, Y_2, Y_1, N, N_{[n-1]}, \ldots, N_{[2]}, N_{[1]}),$$
(B80)

$$A = (-1, 0, 0, \ldots, 0, 0, 0, \ldots, 0),$$
(B81)

$$\Phi_1(X) = \mathrm{diag}(\mathrm{Tr}_{B_n}[N] - N_{[n-1]} \otimes I_{A_n}, \mathrm{Tr}_{B_{n-1}}[N_{[n-1]}]$$
$$- N_{[n-2]} \otimes I_{A_{n-1}}, \ldots, \mathrm{Tr}_{B_2}[N_{[2]}]$$
$$- N_{[1]} \otimes I_{A_2}, \mathrm{Tr}_{B_1}[N_{[1]}]),$$
(B82)

$$B_1 = \mathrm{diag}(0, 0, \ldots, 0, I_{A_1}),$$
(B83)

$$\Phi_2(X) = \mathrm{diag}(N - \lambda \Gamma^{\mathcal{M}^{(n)}}, -Y_n - N, \mathrm{Tr}_{B_n}[Y_n] - Y_{n-1} \otimes I_{A_n},$$
$$\mathrm{Tr}_{B_{n-1}}[Y_{n-1}] - Y_{n-2} \otimes I_{A_{n-1}}, \ldots, \mathrm{Tr}_{B_2}[Y_2]$$
$$- Y_1 \otimes I_{A_2}, \mathrm{Tr}_{B_1}[Y_1]),$$
(B84)

$$B_2 = \mathrm{diag}(0, -\Gamma^{\mathcal{N}^{(n)}}, 0, \ldots, 0, \varepsilon I_{A_1}).$$
(B85)

We should now figure out the adjoints of Φ_1 and Φ_2. Consider that

$$\mathrm{Tr}[Y_i \Phi_i(X)] = \mathrm{Tr}[\Phi_i^\dagger(Y_i)X] \text{ for } i \in \{1, 2\}.$$
(B86)

Set

$$Y_1 = \mathrm{diag}(Z_n, Z_{n-1}, \ldots, Z_2, Z_1).$$
(B87)

Then consider that

$$\mathrm{Tr}[Y_1 \Phi_1(X)] = \mathrm{Tr}[Z_n(\mathrm{Tr}_{B_n}[N] - N_{[n-1]} \otimes I_{A_n})] + \mathrm{Tr}[Z_{n-1}(\mathrm{Tr}_{B_{n-1}}[N_{[n-1]}] - N_{[n-2]} \otimes I_{A_{n-1}})] + \cdots$$
$$+ \mathrm{Tr}[Z_2(\mathrm{Tr}_{B_2}[N_{[2]}] - N_{[1]} \otimes I_{A_2})] + \mathrm{Tr}[Z_1 \mathrm{Tr}_{B_1}[N_{[1]}]]$$
(B88)

$$= \mathrm{Tr}[(Z_n \otimes I_{B_n})N] - \mathrm{Tr}[\mathrm{Tr}_{A_n}[Z_n]N_{[n-1]}] + \mathrm{Tr}[(Z_{n-1} \otimes I_{B_{n-1}})N_{[n-1]}] - \mathrm{Tr}[\mathrm{Tr}_{A_{n-1}}[Z_{n-1}]N_{[n-2]}] + \cdots$$
$$+ \mathrm{Tr}[(Z_2 \otimes I_{B_2})N_{[2]}] - \mathrm{Tr}[\mathrm{Tr}_{A_2}[Z_2]N_{[1]}] + \mathrm{Tr}[(Z_1 \otimes I_{B_1})N_{[1]}]$$
(B89)

$$= \mathrm{Tr}[(Z_n \otimes I_{B_n})N] + \mathrm{Tr}[(Z_{n-1} \otimes I_{B_{n-1}} - \mathrm{Tr}_{A_n}[Z_n])N_{[n-1]}] + \mathrm{Tr}[(Z_{n-2} \otimes I_{B_{n-2}} - \mathrm{Tr}_{A_{n-1}}[Z_{n-1}])N_{[n-2]}] + \cdots$$
$$+ \mathrm{Tr}[(Z_2 \otimes I_{B_2} - \mathrm{Tr}_{A_3}[Z_3])N_{[2]}] + \mathrm{Tr}[(Z_1 \otimes I_{B_1} - \mathrm{Tr}_{A_2}[Z_2])N_{[1]}].$$
(B90)

So then the adjoint of Φ_1 is given by

$$\Phi_1^\dagger(Y_1) = \mathrm{diag}(0, 0, 0, \ldots, 0, 0, Z_n \otimes I_{B_n}, Z_{n-1} \otimes I_{B_{n-1}} - \mathrm{Tr}_{A_n}[Z_n], \ldots, Z_2 \otimes I_{B_2} - \mathrm{Tr}_{A_3}[Z_3], Z_1 \otimes I_{B_1} - \mathrm{Tr}_{A_2}[Z_2]).$$
(B91)

Set

$$Y_2 = \mathrm{diag}(W_{n+2}, W_{n+1}, W_n, W_{n-1}, \ldots, W_2, W_1).$$
(B92)

Then we find that

$$\mathrm{Tr}[Y_2 \Phi_2(X)] = \mathrm{Tr}[W_{n+2}(N - \lambda \Gamma^{\mathcal{M}^{(n)}})] - \mathrm{Tr}[W_{n+1}(Y_n + N)] + \mathrm{Tr}[W_n(\mathrm{Tr}_{B_n}[Y_n] - Y_{n-1} \otimes I_{A_n})]$$
$$+ \mathrm{Tr}[W_{n-1}(\mathrm{Tr}_{B_{n-1}}[Y_{n-1}] - Y_{n-2} \otimes I_{A_{n-1}})] + \cdots + \mathrm{Tr}[W_2(\mathrm{Tr}_{B_2}[Y_2] - Y_1 \otimes I_{A_2})] + \mathrm{Tr}[W_1 \mathrm{Tr}_{B_1}[Y_1]]$$
(B93)

$$= \mathrm{Tr}[W_{n+2}N] - \lambda \mathrm{Tr}[W_{n+2}\Gamma^{\mathcal{M}^{(n)}}] - \mathrm{Tr}[W_{n+1}(Y_n + N)] + \mathrm{Tr}[(W_n \otimes I_{B_n})Y_n] - \mathrm{Tr}[\mathrm{Tr}_{A_n}[W_n]Y_{n-1}]$$
$$+ \mathrm{Tr}[(W_{n-1} \otimes I_{B_{n-1}})Y_{n-1}] - \mathrm{Tr}[\mathrm{Tr}_{A_{n-1}}[W_{n-1}]Y_{n-2}] + \cdots + \mathrm{Tr}[(W_2 \otimes I_{B_2})Y_2] - \mathrm{Tr}[\mathrm{Tr}_{A_2}[W_2]Y_1]$$
$$+ \mathrm{Tr}[(W_1 \otimes I_{B_1})Y_1]$$
(B94)

$$= \mathrm{Tr}[(W_{n+2} - W_{n+1})N] - \lambda \mathrm{Tr}[W_{n+2}\Gamma^{\mathcal{M}^{(n)}}] + \mathrm{Tr}[(W_n \otimes I_{B_n} - W_{n+1})Y_n] + \mathrm{Tr}[(W_{n-1} \otimes I_{B_{n-1}} - \mathrm{Tr}_{A_n}[W_n])Y_{n-1}]$$
$$+ \mathrm{Tr}[(W_{n-2} \otimes I_{B_{n-2}} - \mathrm{Tr}_{A_{n-1}}[W_{n-1}])Y_{n-2}] + \cdots + \mathrm{Tr}[(W_2 \otimes I_{B_2} - \mathrm{Tr}_{A_3}[W_3])Y_2]$$
$$+ \mathrm{Tr}[(W_1 \otimes I_{B_1} - \mathrm{Tr}_{A_2}[W_2])Y_1].$$
(B95)

So then the adjoint of Φ_2 is given by

$$\Phi_2^\dagger(Y_2) = \mathrm{diag}(-\mathrm{Tr}[W_{n+2}\Gamma^{\mathcal{M}^{(n)}}], W_n \otimes I_{B_n} - W_{n+1}, W_{n-1} \otimes I_{B_{n-1}} - \mathrm{Tr}_{A_n}[W_n], \ldots, W_2 \otimes I_{B_2} - \mathrm{Tr}_{A_3}[W_3],$$
$$W_1 \otimes I_{B_1} - \mathrm{Tr}_{A_2}[W_2], W_{n+2} - W_{n+1}, 0, \ldots, 0, 0).$$
(B96)

Adding $\Phi_1^\dagger(Y_1)$ and $\Phi^\dagger(Y_2)$ gives

$$\Phi_1^\dagger(Y_1) + \Phi^\dagger(Y_2) = \mathrm{diag}(-\mathrm{Tr}[W_{n+2}\Gamma^{\mathcal{M}^{(n)}}], W_n \otimes I_{B_n} - W_{n+1}, W_{n-1} \otimes I_{B_{n-1}} - \mathrm{Tr}_{A_n}[W_n], \ldots, W_2 \otimes I_{B_2} - \mathrm{Tr}_{A_3}[W_3],$$

$$W_1 \otimes I_{B_1} - \mathrm{Tr}_{A_2}[W_2], W_{n+2} - W_{n+1} + Z_n \otimes I_{B_n}, Z_{n-1} \otimes I_{B_{n-1}}$$

$$- \mathrm{Tr}_{A_n}[Z_n], \ldots, Z_2 \otimes I_{B_2} - \mathrm{Tr}_{A_3}[Z_3], Z_1 \otimes I_{B_1} - \mathrm{Tr}_{A_2}[Z_2]). \tag{B97}$$

Then the inequality $\Phi_1^\dagger(Y_1) + \Phi_2^\dagger(Y_2) \geqslant A$ is equivalent to the following set of inequalities:

$$-\mathrm{Tr}[W_{n+2}\Gamma^{\mathcal{M}^{(n)}}] \geqslant -1, \tag{B98}$$

$$W_n \otimes I_{B_n} - W_{n+1} \geqslant 0, \tag{B99}$$

$$W_{n-1} \otimes I_{B_{n-1}} - \mathrm{Tr}_{A_n}[W_n] \geqslant 0, \tag{B100}$$

$$\vdots$$

$$W_2 \otimes I_{B_2} - \mathrm{Tr}_{A_3}[W_3] \geqslant 0, \tag{B101}$$

$$W_1 \otimes I_{B_1} - \mathrm{Tr}_{A_2}[W_2] \geqslant 0, \tag{B102}$$

$$W_{n+2} - W_{n+1} + Z_n \otimes I_{B_n} \geqslant 0, \tag{B103}$$

$$Z_{n-1} \otimes I_{B_{n-1}} - \mathrm{Tr}_{A_n}[Z_n] \geqslant 0, \tag{B104}$$

$$\vdots$$

$$Z_2 \otimes I_{B_2} - \mathrm{Tr}_{A_3}[Z_3] \geqslant 0, \tag{B105}$$

$$Z_1 \otimes I_{B_1} - \mathrm{Tr}_{A_2}[Z_2] \geqslant 0, \tag{B106}$$

which can be rewritten as

$$\mathrm{Tr}[W_{n+2}\Gamma^{\mathcal{M}^{(n)}}] \leqslant 1, \tag{B107}$$

$$W_n \otimes I_{B_n} \geqslant W_{n+1}, \tag{B108}$$

$$W_{n-1} \otimes I_{B_{n-1}} \geqslant \mathrm{Tr}_{A_n}[W_n], \tag{B109}$$

$$\vdots$$

$$W_2 \otimes I_{B_2} \geqslant \mathrm{Tr}_{A_3}[W_3], \tag{B110}$$

$$W_1 \otimes I_{B_1} \geqslant \mathrm{Tr}_{A_2}[W_2], \tag{B111}$$

$$W_{n+2} + Z_n \otimes I_{B_n} \geqslant W_{n+1}, \tag{B112}$$

$$Z_{n-1} \otimes I_{B_{n-1}} \geqslant \mathrm{Tr}_{A_n}[Z_n], \tag{B113}$$

$$\vdots$$

$$Z_2 \otimes I_{B_2} \geqslant \mathrm{Tr}_{A_3}[Z_3], \tag{B114}$$

$$Z_1 \otimes I_{B_1} \geqslant \mathrm{Tr}_{A_2}[Z_2]. \tag{B115}$$

The dual objective function is given by

$$-\mathrm{Tr}[B_1 Y_1] - \mathrm{Tr}[B_2 Y_2]$$

$$= -\mathrm{Tr}[Z_1] + \mathrm{Tr}[\Gamma^{\mathcal{N}^{(n)}} W_{n+1}] - \varepsilon \mathrm{Tr}[W_1]. \tag{B116}$$

So then the dual can be written as

$$\sup_{\substack{Z_n, \ldots, Z_1 \in \mathrm{Herm}, \\ W_{n+2}, \ldots, W_1 \geqslant 0}} -\mathrm{Tr}[Z_1] + \mathrm{Tr}[\Gamma^{\mathcal{N}^{(n)}} W_{n+1}] - \varepsilon \mathrm{Tr}[W_1]$$

$$\tag{B117}$$

subject to

$$\mathrm{Tr}[W_{n+2}\Gamma^{\mathcal{M}^{(n)}}] \leqslant 1, \tag{B118}$$

$$W_n \otimes I_{B_n} \geqslant W_{n+1}, \tag{B119}$$

$$W_{n-1} \otimes I_{B_{n-1}} \geqslant \mathrm{Tr}_{A_n}[W_n], \tag{B120}$$

$$\vdots$$

$$W_2 \otimes I_{B_2} \geqslant \mathrm{Tr}_{A_3}[W_3], \tag{B121}$$

$$W_1 \otimes I_{B_1} \geqslant \mathrm{Tr}_{A_2}[W_2], \tag{B122}$$

$$W_{n+2} + Z_n \otimes I_{B_n} \geqslant W_{n+1}, \tag{B123}$$

$$Z_{n-1} \otimes I_{B_{n-1}} \geqslant \mathrm{Tr}_{A_n}[Z_n], \tag{B124}$$

$$\vdots$$

$$Z_2 \otimes I_{B_2} \geqslant \mathrm{Tr}_{A_3}[Z_3], \tag{B125}$$

$$Z_1 \otimes I_{B_1} \geqslant \mathrm{Tr}_{A_2}[Z_2]. \tag{B126}$$

Since all of the Z variables are Hermitian, we can make the substitution $Z_i \to -Z_i$ without affecting the optimal value. The final form is then

$$\sup_{\substack{Z_n, \ldots, Z_1 \in \mathrm{Herm}, \\ W_{n+2}, \ldots, W_1 \geqslant 0}} \mathrm{Tr}[Z_1] + \mathrm{Tr}[\Gamma^{\mathcal{N}^{(n)}} W_{n+1}] - \varepsilon \mathrm{Tr}[W_1] \tag{B127}$$

subject to

$$\mathrm{Tr}[W_{n+2}\Gamma^{\mathcal{M}^{(n)}}] \leqslant 1, \tag{B128}$$

$$W_n \otimes I_{B_n} \geqslant W_{n+1}, \tag{B129}$$

$$W_{n-1} \otimes I_{B_{n-1}} \geqslant \mathrm{Tr}_{A_n}[W_n], \tag{B130}$$

$$\vdots$$

$$W_2 \otimes I_{B_2} \geqslant \mathrm{Tr}_{A_3}[W_3], \tag{B131}$$

$$W_1 \otimes I_{B_1} \geqslant \mathrm{Tr}_{A_2}[W_2], \tag{B132}$$

$$W_{n+2} \geqslant W_{n+1} + Z_n \otimes I_{B_n}, \tag{B133}$$

$$\mathrm{Tr}_{A_n}[Z_n] \geqslant Z_{n-1} \otimes I_{B_{n-1}}, \tag{B134}$$

$$\vdots$$

$$\mathrm{Tr}_{A_3}[Z_3] \geqslant Z_2 \otimes I_{B_2}, \tag{B135}$$

$$\mathrm{Tr}_{A_2}[Z_2] \geqslant Z_1 \otimes I_{B_1}. \tag{B136}$$

[1] C. W. Helstrom, Quantum detection and estimation theory, J. Stat. Phys. **1**, 231 (1969).

[2] A. S. Holevo, Statistical decision theory for quantum systems, J. Multivariate Anal. **3**, 337 (1973).

[3] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).

[4] F. Hiai and D. Petz, The proper formula for relative entropy and its asymptotics in quantum probability, Commun. Math. Phys. **143**, 99 (1991).

[5] T. Ogawa and H. Nagaoka, Strong converse and Stein's lemma in quantum hypothesis testing, IEEE Trans. Inf. Theory **46**, 2428 (2000).

[6] A. Yu Kitaev, Quantum computations: Algorithms and error correction, Russ. Math. Surv. **52**, 1191 (1997).

[7] D. Aharonov, A. Kitaev, and N. Nisan, Quantum Circuits with Mixed States, in *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing* (ACM, New York, 1998), pp. 20–30.

[8] A. M. Childs, J. Preskill, and J. Renes, Quantum information and precision measurement, J. Mod. Opt. **47**, 155 (2000).

[9] A. Acin, Statistical Distinguishability Between Unitary Operations, Phys. Rev. Lett. **87**, 177901 (2001).

[10] B. Rosgen and J. Watrous, On the hardness of distinguishing mixed-state quantum computations, in *Proceedings of the 20th IEEE Conference on Computational Complexity*, June 2005 (IEEE, New York, 2005), pp. 344–354,.

[11] A. Gilchrist, N. K. Langford, and M. A. Nielsen, Distance measures to compare real and ideal quantum processes, Phys. Rev. A **71**, 062310 (2005).

[12] K. Matsumoto, Reverse test and characterization of quantum relative entropy, arXiv:1010.1030.

[13] X. Wang and M. M. Wilde, Resource theory of asymmetric distinguishability, Phys. Rev. Research **1**, 033170 (2019).

[14] X. Wang and M. M. Wilde, Resource theory of asymmetric distinguishability for quantum channels, Phys. Rev. Research **1**, 033169 (2019).

[15] G. Gutoski and J. Watrous, Toward a general theory of quantum games, in *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing - STOC '07* (ACM, New York, 2007), p. 565.

[16] G. Gutoski, Quantum strategies and local operations, Ph.D. thesis, 2009, https//uwspace.uwaterloo.ca/handle/10012/4903.

[17] G. Gutoski, On a measure of distance for quantum strategies, J. Math. Phys. **53**, 032202 (2012).

[18] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Transforming quantum operations: Quantum supermaps, Europhys. Lett. **83**, 30004 (2008).

[19] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Memory Effects in Quantum Channel Discrimination, Phys. Rev. Lett. **101**, 180501 (2008).

[20] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Theoretical framework for quantum networks, Phys. Rev. A **80**, 022339 (2009).

[21] M. Hayashi, Discrimination of two channels by adaptive methods and its application to quantum system, IEEE Trans. Inf. Theory **55**, 3807 (2009).

[22] M. M. Wilde, M. Berta, C. Hirche, and E. Kaur, Amortized channel divergence for asymptotic quantum channel discrimination, Lett. Math. Phys. **110**, 2277 (2020).

[23] K. Fang, O. Fawzi, R. Renner, and D. Sutter, A Chain Rule for the Quantum Relative Entropy, Phys. Rev. Lett. **124**, 100501 (2020).

[24] A. W. Harrow, A. Hassidim, D. W. Leung, and J. Watrous, Adaptive versus non-adaptive strategies for quantum channel discrimination, Phys. Rev. A **81**, 032339 (2010).

[25] G. Gour, Comparison of quantum channels by superchannels, IEEE Trans. Inf. Theory **65**, 5880 (2019).

[26] G. Gutoski, A. Rosmanis, and J. Sikora, Fidelity of quantum strategies with applications to cryptography, Quantum **2**, 89 (2018).

[27] G. Chiribella and D. Ebler, Optimal quantum networks and one-shot entropies, New J. Phys. **18**, 093053 (2016).

[28] S. Rethinasamy and M. M. Wilde, Relative entropy and catalytic relative majorization, Phys. Rev. Research **2**, 033455 (2020).

[29] F. Buscemi and N. Datta, The quantum capacity of channels with arbitrarily correlated noise, IEEE Trans. Inf. Theory **56** (3), 1447 (2010).

[30] F. G. S. L. Brandao and N. Datta, One-shot rates for entanglement manipulation under non-entangling maps, IEEE Trans. Inf. Theory **57**, 1754 (2011).

[31] L. Wang and R. Renner, One-Shot Classical-Quantum Capacity and Hypothesis Testing, Phys. Rev. Lett. **108**, 200501 (2012).

[32] N. Datta, Min- and max- relative entropies and a new entanglement monotone, IEEE Trans. Inf. Theory **55**, 2816 (2009).

[33] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, England, 2010).

[34] L. Chirolli and G. Burkard, Decoherence in solid state qubits, Adv. Phys. **57**, 225 (2008).

[35] J. Bavaresco, M. Murao, and M. T. Quintino, Strict hierarchy between parallel, sequential, and indefinite-causal-order strategies for channel discrimination, arXiv:2011.08300.

[36] F. Salek, M. Hayashi, and A. Winter, When are adaptive strategies in asymptotic quantum channel discrimination useful? arXiv:2011.06569.

[37] M. Rexiti and S. Mancini, Discriminating qubit amplitude damping channels, J. Phys. A: Math. Theor. **54**, 165303 (2021).

[38] J. L. Pereira and S. Pirandola, Bounds on amplitude-damping-channel discrimination, Phys. Rev. A **103**, 022610 (2021).