

# Distinguishability, Ensemble Steering, and the No-Signaling Principle

Joonwoo Bae

*Center for Quantum Technologies, National University of Singapore, Singapore 117543,  
ICFO–Institut de Ciències Fotòniques, 08860 Castelldefels (Barcelona), Spain.*

We consider a fundamental operational task: distinguishing systems in different states in the framework of generalized probabilistic theories. We provide a general formalism of minimum-error discrimination of states in convex optimization. With the formalism established, we show that the distinguishability is generally a global property assigned to the ensemble of given states rather than other details of a given state space or pairwise relations of given states. Then, we consider bipartite systems where ensemble steering is possible, and show that with two operational tasks (ensemble steering and satisfying the no-signaling condition) the distinguishability is tightly determined. The result is independent of the structure of the state space. This concludes that the distinguishability is generally determined by the compatibility between two tasks, ensemble steering of states and the no-signaling principle.

## 1 Introduction

One of the most fundamental tasks of information processing is to distinguish different signals that deliver distinct messages. If quantum systems are used to carry messages from one party to another, we are led to the problem of discriminating between different quantum states. In a naive way, if messages are encoded in distinguishable quantum states i.e. orthogonal states, which are then transmitted through a noisy channel, during which the quantum states may interact with the environment resulting in a state that may no longer be orthogonal. Non-orthogonal quantum states cannot be perfectly distinguished in a single-shot manner, that is, they are indistinguishable. Then, for the practical information application of quantum systems, one has to devise methods of quantum state discrimination and optimization of measurement for these purposes.

Interestingly, the indistinguishability of non-orthogonal quantum states seems to have a deeper meaning than merely a problem that can be considered for practical applications. While indistinguishability itself has appeared in contexts of classical information processing through probabilistic systems which cannot be distinguished perfectly, recent works have shown that it cannot be consistently reproduced by probabilistic classical systems only [1]. Indistinguishability of quantum states has been indeed a fundamental reason behind many quantum information applications. For information-theoretic purposes, it is useful to have a direct characterization of the distinguishability in an operational way, however, there is little known so far in this direction.

In this work, we provide an operational characterization of the indistinguishability of quantum states in terms of two operational tasks, ensemble steering and the no-signaling principle. We show that a physical theory in general, where a system is described as a box producing probabilistic outcomes, indistinguishability is tightly connected to the competition between two tasks: one is ensemble steering on sources i.e. the ability of manipulating states of systems from a distance in a remote way before

measurement happens on systems; and the other is the no-signaling constraint on measurement outcomes. In other words, distinguishability cannot be improved even better if ensemble steering is allowed and the no-signaling principle is respected. Then, quantum theory is merely an instance of the case; the quantum indistinguishability can be characterized by ensemble steering on quantum states and the no-signaling principle on measurement outcomes.

## 2 Preliminaries

Before we proceed to the main and technical details, let us set notations and basic definitions. We begin with the framework in GPTs where notions of states, measurement, and their relations are generalized [2, 3, 4]. Here we will make use of the mathematical framework in Ref. [3].

In a GPT, the set of states, denoted by  $\Omega$ , consists of all possible states that a system can be prepared in. Any probabilistic mixture of states, i.e.  $pw_1 + (1-p)w_2 \in \Omega$  for  $w_1, w_2 \in \Omega$  and probability  $p$  is also a state, i.e.  $\Omega$  is convex. A general mapping from states to probabilities is called *effects* and described by linear functionals  $\Omega \rightarrow [0, 1]$ . A measurement denoted by  $s$  corresponds to set of effects,  $E^{(s)} = \{e_x^{(s)}\}_{x=1}^N$ , by which the probability of getting outcome  $x$  given a state  $w$  is,  $p(x|s) = e_x^{(s)}[w]$ . A unit effect  $u$  means a measurement that occurs, that is,  $u[w] = 1$  for all  $w \in \Omega$ , so that for any measurement  $s$ , it holds  $\sum_x e_x^{(s)} = u$ . As effects are dual to the state space, they are also convex.

Here, we identify state distinguishability with minimal error, or equivalently, the maximal success probability in a guessing task for different states. This is called minimum-error state discrimination, and can be described by a game with two parties, Alice and Bob, as follows. Suppose there is a set of states agreed by them in advance, and Alice prepares a system in one of  $N$  states with some probability and gives it to Bob. If Bob makes a correct guess, their score is given 1, otherwise 0. Their goal is to maximize the average score over all measurements. We write the states by  $\{w_x\}_{x=1}^N$  and Alice's *a priori* probabilities by  $\{q_x\}_{x=1}^N$ , and then by  $\{q_x, w_x\}_{x=1}^N$  altogether. Bob has to find optimal measurement to maximize the score. We write by  $p_{B|A}(x|y) = e_x[w_y]$  the probability that Bob makes a guess  $w_x$  when state  $w_y$  is given by Alice. Now, the goal is to find the maximal success probability of making a correct guess, the so-called *guessing probability* in the following,

$$p_{\text{guess}} := \max \sum_{x=1}^N q_x p_{B|A}(x|x) = \max_{\{e_x\}_{x=1}^N} \sum_{x=1}^N q_x e_x[w_x] \quad (1)$$

where the maximization runs over all effects. Note that GPTs are generally not self-dual, meaning an isomorphism between two spaces does not exist in general [5].

## 3 Indistinguishability via optimal discrimination

Since state and effect spaces are convex, it is possible to formalize the optimization task in Eq. (1) in the convex optimization framework [6]. For states  $\{q_x, w_x\}_{x=1}^N$ , we take the form in Eq. (1) as the primal problem denoted by  $p^*$

$$p^* = \max \left\{ \sum_{x=1}^N q_x e_x[w_x] \mid e_x \geq 0 \forall x, \sum_{x=1}^N e_x = u \right\} \quad (2)$$

and derive its dual  $d^*$ , as follows

$$d^* = \min \{ u[K] \mid K \geq q_x w_x, x = 1, \dots, N \} \quad (3)$$

where inequalities mean the order relation in the convex set: by  $e_x \geq 0$ , it is meant that  $e_x[w] \geq 0$  for all  $w \in \Omega$ , and by  $K \geq q_x w_x$ , that  $e[K - q_x w_x] \geq 0$  for all effects  $e$ . Following from the so-called Slater's constraint quantification in convex optimization, a sufficient condition for the strong duality is the strict feasibility, which is the existence of a strictly feasible point of parameters. For instance, primal parameters  $\{e_x = u/N\}_{x=1}^N$  are in the case, since  $e_x[w_y] > 0 \forall x, y$  and  $\sum_x e_x = 1$ . Thus, the strong duality holds true, meaning that both solutions from primal and dual problems are equal, i.e.  $p^* = d^*$ .

There is another approach called the *complementarity problem* that generalizes convex optimization, in which the optimality conditions of a given optimization problem are directly analyzed. This approach has been applied to quantum state discrimination in the context of two-state discrimination in Refs. [7, 8]. The approach has been cited and generalized to GPTs in Ref. [9] under the assumption that conjugate states, equivalent to complementary states in the below, do exist. There those classes of quantum states for which conjugate states exist are named the Helstrom family, the existence of which has remained in question. Therefore, the results are valid only under the unproven assumption. In fact, a formal statement of the complementary problem for optimal state discrimination in GPTs, that is, a formal way that provides solutions in an explicit way, has been shown in Ref. [10]. In what follows, we provide the complementary problem in a formal way via convex optimization, through which the existence of all parameters is immediately guaranteed.

The complementarity problem deals with both primal and dual parameters in Eqs. (2) and (3) which in general is not more efficient than the convex optimization approach. The advantage, however, lies at the fact that generic structures existing in the problem are exploited. The optimality conditions can be summarized by the so-called Karush-Kuhn-Tucker (KKT) conditions, which are constraints listed in Eqs. (2) and (3), together with the followings,

$$\text{(Symmetry parameter)} \quad K = q_x w_x + r_x d_x, \quad \forall x \quad (4)$$

$$\text{(Orthogonality)} \quad e_x[r_x d_x] = 0, \quad \forall x, \quad (5)$$

where  $r_x \in [0, 1]$  for all  $x$ , and  $\{d_x\}_{x=1}^N$  which we call complementary states are normalized, i.e.  $u[d_x] = 1$ . The first condition, the symmetry parameter, follows from the Lagrangian stability and shows that for any discrimination problem e.g.  $\{q_x, w_x\}_{x=1}^N$ , there exists a single parameter  $K$  which is decomposed into  $N$  different ways with given states and complementary states  $\{r_x, d_x\}_{x=1}^N$ . Then, the second condition in Eq. (5) from the complementary slackness characterizes optimal effects by the orthogonality relation between complementary states and optimal effects. These generalize optimality conditions for quantum states in Refs. [20] [21] to all GPTs, see also different forms of optimality conditions [?]. Primal and dual parameters satisfying the KKT conditions are automatically optimal parameters that provide solutions to optimal discrimination. Moreover, for the problem here, recall that the strong duality holds i.e.  $p^* = d^*$ . Conversely, the fact that the strong duality holds in Eqs. (2) and (3) implies the existence of optimal parameters which satisfy KKT conditions and give the guessing probability in Eq. (1). Note that we derive all these from the fact that state spaces of GPTs are convex.

The complementarity problem can then be formalized in a geometric way. We first remark that, in optimality conditions in Eqs. (4) and (5), constraints for states and effects are separated. The symmetry parameter  $K$  is characterized on a state space and gives the guessing probability, see Eq. (3), i.e.  $p_{\text{guess}} = u[K] = q_x + r_x$ . Then, the guessing probability can be found by searching complementary states  $\{r_x, d_x\}_{x=1}^N$  fulfilling Eq. (4) on the state space. This can be described in a systematic way, as follows. Let us define a polytope denoted by  $\mathcal{P}(\{q_x, w_x\}_{x=1}^N)$  of given states in the state space: each vertex of the polytope corresponds to unnormalized state  $q_x w_x$  for  $x = 1, \dots, N$ . Then, the polytope of complementary states,  $\mathcal{P}(\{r_x, d_x\}_{x=1}^N)$ , is in fact immediately congruent to  $\mathcal{P}(\{q_x, w_x\}_{x=1}^N)$  in the state space, since the

following holds from Eq. (4),

$$q_x w_x - q_y w_y = r_y d_y - r_x d_x, \text{ for all } x, y, \quad (6)$$

which shows that corresponding lines of two polytopes  $\mathcal{P}(\{q_x, w_x\}_{x=1}^N)$  and  $\mathcal{P}(\{r_x, d_x\}_{x=1}^N)$  are of equal lengths and anti-parallel. Then, from the underlying geometry of the state space, one can find complementary states by putting two congruent polytopes such that the condition in Eq. (4) holds. Optimal effects can be found from the orthogonal relation in Eq. (5), accordingly.

For *a priori* probabilities given as  $q_x = 1/N$ , the guessing probability becomes even simpler. First, it follows  $r_x = r_y$  for all  $x, y$ : this is obtained from the expression  $p_{\text{guess}} = q_x + r_x$  for any  $x$ , see Eqs. (4) and (3). Denoted by  $r := r_x$  for all  $x$ , the guessing probability is now,

$$p_{\text{guess}} = \frac{1}{N} + r, \text{ with } r = \frac{\|\frac{1}{N}w_x - \frac{1}{N}w_y\|}{\|d_x - d_y\|} \quad (7)$$

where the expression of  $r$  follows from the condition in Eq. (6) with a distance measure  $\|\cdot\|$  that can be defined in the state space. The parameter  $r$  has a meaning as the ratio between two polytopes,  $\mathcal{P}(\{1/N, w_x\}_{x=1}^N)$  of given states, and  $\mathcal{P}(\{d_x\}_{x=1}^N)$  of only complementary states.

## 4 Tightness

We now move to a bipartite scenario where the state space of two parties is constrained such that ensemble steering is possible. This means that Alice can steer to any decomposition of Bob's ensemble. In quantum theory, the notion of steering was introduced by Schrödinger [11] and then, with specification to a bipartite Hilbert space, formalized by the so-called Gisin-Hughston-Jozsa-Wootters theorem [12, 13]. Note that ensemble steering does not yet single out quantum theory among GPTs [14]. We also distinguish the extension from the purification lemma which fully characterizes quantum theory [15].

In what follows, we apply the theoretical tools developed so far, and show that for any GPTs endowed with ensemble steering, the distinguishability is immediately determined by a way of excluding instantaneous communication. We first derive a bound to the optimal distinguishability in a given GPT by the no-signaling condition, and then we prove that the bound is tight, i.e. it can be achieved within the given GPT. The result is independent of particular properties of a state space.

Let us incorporate state discrimination to the following non-signaling framework. Let  $\{q_x, w_x\}_{x=1}^N$  denote the states among which we are interested in discriminating. Suppose Alice steers the ensemble of Bob, denoted by  $w_B$ , in  $N$  different decompositions. That is, the ensemble has  $N$  different decomposition as  $w_B = w_B^{(x)}$  for  $x = 1, \dots, N$  where

$$w_B^{(x)} = p_x w_x + (1 - p_x) c_x, \text{ with } q_x = \frac{p_x}{\sum_{x'=1}^N p_{x'}} \quad (8)$$

with some states  $\{c_x\}_{x=1}^N$  and probabilities  $\{p_x\}_{x=1}^N$ . By ensemble steering, we mean that any of the  $N$  decompositions of Bob's ensemble can be prepared by Alice's steering. Since Bob holds an identical ensemble, his measurement gains no knowledge about which decomposition is given, until Alice announces about her steering. The non-signaling condition is thus naturally imposed.

The distinguishability on  $\{q_x, w_x\}_{x=1}^N$  is then constrained by the non-signaling condition as follows. Assume that Bob optimizes measurement to guess which state among  $\{w_x\}_{x=1}^N$  exists in his ensemble. The strategy is, once the state  $w_x$  is found, he concludes his ensemble is in the decomposition  $w_B^{(x)}$ ,

see Eq. (8), by which he also guesses Alice's steering strategy. Then, by the no-signaling condition, discrimination among states  $\{w_x\}_{x=1}^N$  must be constrained so that Bob would not learn Alice's steering better than the random guess.

We now derive an upper bound to the guessing probability by the no-signaling condition. Let  $P_{B|A}(x|y)$  denote the probability that, while Alice has actually steered ensemble  $w_B^{(y)}$ , Bob concludes his ensemble in  $w_B^{(x)}$  by discriminating among  $\{q_x, w_x\}_{x=1}^N$ . The no-signaling condition [4] [16] implies the following constraint

$$\sum_{x=1}^N P_{B|A}(x|x) \leq 1. \quad (9)$$

If the condition is not fulfilled, one can explicitly construct a superluminal communication protocol [17]. Then, recall Bob's strategy of guessing Alice's steering: to guess Alice's steering strategy, he attempts to distinguish ensemble decompositions  $\{w_B^{(x)}\}_{x=1}^N$  by exploiting optimal discrimination of states  $\{w_x\}_{x=1}^N$  existing in the ensemble.

If Alice has steered Bob's ensemble  $w_B^{(x)}$ , Bob's correct conclusion happens when i)  $w_x$  is given, which appears with probability  $p_x$ , and ii) measurement gives a correct answer, that is, with probability  $p_x P_{B|A}(x|x)$ . In the strategy, there can be contribution in measurement from the other state  $c_x$  in the ensemble with probability  $1 - p_x$ . Thus, it holds,  $p_x P_{B|A}(x|x) \leq P_{B|A}(x|x)$ . In addition, recall that measurement is optimized for discrimination among  $\{q_x, w_x\}_{x=1}^N$ , since the *a priori* probability for state  $w_x$  among  $\{w_x\}_{x=1}^N$  is given by  $q_x$ , see Eq. (8). From the no-signaling condition in Eq. (9), we have  $\sum_{x=1}^N p_x P_{B|A}(x|x) \leq 1$ , from which we have

$$p_{\text{guess}} = \max_x \sum_x q_x P_{B|A}(x|x) \leq \frac{1}{p_1 + \dots + p_N}. \quad (10)$$

Thus, an upper bound to the distinguishability is obtained from the no-signaling condition, and expressed in terms of parameters  $\{p_x\}_{x=1}^N$  of steering each state in  $\{w_x\}_{x=1}^N$ .

We then show that the bound is indeed tight, i.e. it can be achieved within a given GPT. We show the tightness by proving that, for any set  $\{q_x, w_x\}_{x=1}^N$ , the optimal discrimination characterized by the KKT conditions implies the existence of both an identical ensemble in Eq. (8) and effects achieving the bound in Eq. (10).

Recall the general method of optimal discrimination, the existence of a symmetry parameter  $K$  that completely characterizes the optimal distinguishability, see Eq. (4). The parameter has  $N$  decompositions with complementary states  $\{r_x, d_x\}_{x=1}^N$ . Its normalization  $\tilde{K} = K/u[K]$  shows, for each  $x$ ,

$$\tilde{K} = p_x w_x + (1 - p_x) d_x, \text{ with } p_x = q_x/u[K]. \quad (11)$$

This corresponds to ensemble steering in Eq. (8). Recall the dual problem in Eq. (3) which gives the guessing probability in GPTs, as  $p_{\text{guess}} = u[K]$ . Using the identity  $\sum_{x=1}^N q_x = 1$  and the relation  $p_x = q_x/u[K]$ , the solution in the dual problem can be computed as,  $u(K) = (\sum_{x=1}^N p_x)^{-1}$ . This shows that the bound in Eq. (10) is already achieved within a given GPT, and hence the tightness is shown. In addition, optimal effects also exist with complementary states, see Eq. (5).

## 5 Summary

We have developed and established a general method of distinguishing states in GPTs. This generalizes i) the geometric formulation [18, 19] and ii) optimality conditions [20] [21] in the quantum case to

GPTs. The formulation is also illustrated with an example (the four-state polygon system) a particularly interesting case where the bipartite extension shows the maximally non-local correlations [4] [5]. It is also shown that distinguishability and non-locality are independent resources. We also remark that in GPTs i) measurement for the optimal discrimination is generally not unique, and ii) sometimes no measurement give an optimal strategy, thus similar to the results in quantum cases in [18, 19] [22] [23]. With the general formalism and tools developed, we have shown that for GPTs where ensemble steering is possible, the distinguishability can be determined by no-signaling condition. State distinguishability was also shown to be dictated by the relation between ensemble steering and the no-signaling condition. This also gives a physical motivation for why we have the Born rule in Hilbert space quantum mechanics as opposed to mathematical motivations (e.g. Gleason's theorem [24]): we can have both ensemble steering yet satisfy the no-signalling principle.

## References

- [1] Matthew Leifer & Owen J. E. Maroney (2013): *Maximally epistemic interpretations of the quantum state and contextuality*. *Phys. Rev. Lett* 110(120401), doi:10.1103/PhysRevLett.110.120401.
- [2] George D. Birkhoff & John von Neumann (1936): *The logic of quantum mechanics*. *Annals of Math.* 37, pp. 823–834, doi:10.2307/1968621.
- [3] Howard Barnum, Jonathan Barrett, Matthew Leifer & Alexander Wilce (2007): *Generalized no-broadcasting theorem**Phys. Rev. Lett* 99(240501), doi:10.1103/PhysRevLett.99.240501.
- [4] Jonathan Barrett (2007): *Information processing in generalized probabilistic theories*. *Phys. Rev. A* 75(032304), doi:10.1103/PhysRevA.75.032304.
- [5] Peter Janotta, Christian Gogolin, Jonathan Barrett & Nicolas Brunner (2011): *Limits on nonlocal correlations from the structure of the local state space*. *New J. Phys.* 13(063024), doi:10.1088/1367-2630/13/6/063024.
- [6] Stephen Boyd & Lieven Vandenberghe (2004): *Convex Optimization*. Cambridge University Press, doi:10.1017/cbo9780511804441.
- [7] Won-Young Hwang (2005): *Helstrom theorem from the no-signaling condition*. *Phys. Rev. A* 71(062315), doi:10.1103/PhysRevA.71.062315.
- [8] Joonwoo Bae, Jae-Weon Lee, Jaewan Kim & Won-Young Hwang(2008): *Optimality of minimum-error discrimination by the no-signalling condition*. *Phys. Rev. A* 78(022335), doi:10.1103/PhysRevA.78.022335.
- [9] Gen Kimura, Takayuki Miyadera & Hideki Imai(2009): *Optimal State Discrimination in General Probabilistic Theories*. *Phys. Rev. A* 79(062306), doi:10.1103/PhysRevA.79.062306.
- [10] Joonwoo Bae (2012): *Distinguishability, Ensemble Steering, and the No-Signaling Principle*. <http://arxiv.org/abs/1210.3125v2>.
- [11] Erwin Schrödinger(1935): *Discussion of Probability Relations between Separated Systems*. *Proc. Cambridge Phil. Soc.* 31(555-563), doi:10.1017/S0305004100013554.
- [12] Nicolas Gisin(1989): *Stochastic Quantum Dynamics and Relativity*. *Helv. Phys. Acta* 62(363-371).
- [13] Lane P. Hughston, Richard Jozsa & William K. Wootters(1993): *A complete classification of quantum ensembles having a given density matrix*. *Phys. Lett. A.* 183(14-18), doi:10.1016/0375-9601(93)90880-9.
- [14] Howard Barnum, Carl P. Gaebler & Alexander Wilce(2009): *Ensemble Steering, Weak Self-Duality, and the Structure of Probabilistic Theories*. <http://arxiv.org/abs/0912.5532>.
- [15] Giulio Chiribella, Giacomo Mauro D'Ariano & Paolo Perinotti(2011): *Informational derivation of quantum theory*. *Phys. Rev. A* 84(012311), doi:10.1103/PhysRevA.84.012311.
- [16] Lluís Masanes, Antonio Acín & Nicolas Gisin(2006): *General properties of nonsignaling theories*. *Phys. Rev. A* 73(012112), doi:10.1103/PhysRevA.73.012112.

- [17] Joonwoo Bae, Won-Young Hwang & Yeong-Deok Han(2011): *No-Signaling Principle Can Determine Optimal Quantum State Discrimination*. *Phys. Rev. Lett.* 107(170403), doi:10.1103/PhysRevLett.107.170403.
- [18] Joonwoo Bae(2013): *Structure of minimum-error quantum state discrimination*. *New J. Phys.* 15(073037), doi:10.1088/1367-2630/15/7/073037.
- [19] Joonwoo Bae & Won-Young Hwang(2013): *Minimum-error discrimination of qubit states: Methods, solutions, and properties*. *Phys. Rev. A* 87(012334), doi:10.1103/PhysRevA.87.012334.
- [20] Alexander S. Holevo(1974): *Remarks on optimal measurements*. *Problems of Information Transmission* 10(317-320).
- [21] Horace P. Yuen, Robert S. Kennedy & Melvin Lax(1975): *Optimum testing of multiple hypotheses in quantum detection theory*. *IEEE Trans. Inf. Theory* IT-21(2), doi:10.1103/PhysRevA.73.012112.
- [22] Carl W. Helstrom(1969): *Quantum detection and estimation theory*. *J. of Stat. Phys.* 1(231-252), doi:10.1007/BF01007479.
- [23] Kieran Hunter(2003): *Measurement does not always aid state discrimination*. *Phys. Rev. A* 68(012306).
- [24] Andrew M. Gleason(1957): *Measures on the closed subspaces of a Hilbert space*. *Indiana Univ. Math. J.* 6(4), doi:10.1512/iumj.1957.6.56050.