1. Preliminaries from convex analysis

In the following, we will need some standard terminology and notations from convex analysis. We refer to [1, Secs. 1, 6, 11 and 18] for further details.

We fix a finite dimensional, real and normed linear space V, with dual space V^*. The canonical pairing between an element $v \in V$ and a dual vector $v^* \in V^*$ is denoted by $\langle v , v^* \rangle$. An *affine functional* on V is any map $\Phi : V \to \mathbb{R}$ such that $\Phi(\lambda u + (1-\lambda)v) = \lambda\Phi(u) + (1-\lambda)\Phi(v)$ for all $u, v \in V$ and $\lambda \in \mathbb{R}$. For such a functional, there exist unique $v^* \in V^*$ and $\delta \in \mathbb{R}$ such that

$$\Phi(v) = \langle v , v^* \rangle - \delta =: \Phi[v^*, \delta](v) \qquad \forall v \in V\,. \tag{1}$$

The dual vector v^* as above is the *linear part* of Φ.

Let $\mathscr{C}$ be a convex subset of V. The *affine set* generated by $\mathscr{C}$ is the set $\mathscr{M}(\mathscr{C})$ of all vectors of the form $\lambda_1 x_1 + \ldots + \lambda_n x_n$ for $x_i \in \mathscr{C}$ and $\lambda_i \in \mathbb{R}$ such that $\lambda_1 + \ldots + \lambda_n = 1$ [1, Sec. 1]. There exists a unique vecor subspace $V(\mathscr{C}) \subset V$ (the *translations* of $\mathscr{M}(\mathscr{C})$) such that $\mathscr{M}(\mathscr{C}) = v_0 + V(\mathscr{C})$ for some (hence for any) $v_0 \in \mathscr{M}(\mathscr{C})$. We denote by $V(\mathscr{C})^\perp = \{v^* \in V^* \mid \langle v , v^* \rangle = 0 \ \forall v \in V(\mathscr{C})\}$ the *annihilator subspace* of $V(\mathscr{C})$ in V. The *relative interior* $\mathrm{ri}(\mathscr{C})$ is the set of all interior points of $\mathscr{C}$ with respect to the relative topology of $\mathscr{M}(\mathscr{C})$. An element $z \in \mathscr{C}$ is an *extreme point* of $\mathscr{C}$ if the equality $z = \lambda x + (1-\lambda)y$ with $x, y \in \mathscr{C}$ and $\lambda \in (0,1)$ implies $x = y = z$.

If $\mathscr{C}$ is a compact convex set, its *support function* is the map

$$\delta_{\mathscr{C}} : V^* \to \mathbb{R} \qquad\qquad \delta_{\mathscr{C}}(v^*) = \max\{\langle x , v^* \rangle \mid x \in \mathscr{C}\}\,.$$

Clearly, if $\mathscr{C}' \subseteq \mathscr{C}$ is a compact convex subset, then $\delta_{\mathscr{C}'}(v^*) \leq \delta_{\mathscr{C}}(v^*)$ for all $v^* \in V^*$. The support function satisfies $\delta_{\mathscr{C}}(\alpha v^*) = \alpha\delta_{\mathscr{C}}(v^*)$ for all $\alpha > 0$; moreover, if $v^* - u^* \in V(\mathscr{C})^\perp$, then $\delta_{\mathscr{C}}(v^*) - \delta_{\mathscr{C}}(u^*) = \langle v_0 , v^* - u^* \rangle$ for some (hence for any) $v_0 \in \mathscr{M}(\mathscr{C})$.

2. Certifying a convex property

We are interested in a set of objects $\mathscr{O}$, in which we consider the subset $\mathscr{P} \subset \mathscr{O}$ of all the objects sharing some given property. We assume $\mathscr{P} \neq \emptyset$ and $\mathscr{P} \neq \mathscr{O}$, and we denote by $\mathscr{P}^c = \mathscr{O} \setminus \mathscr{P}$ the subset of all objects which *do not* possess the property; we want to find simple sufficient conditions guaranteeing that an object $x \in \mathscr{O}$ actually belongs to $\mathscr{P}^c$.

In the following, we always suppose that both $\mathscr{P}$ and $\mathscr{O}$ are convex and compact subsets of a finite dimensional, real and normed linear space V. Then, the simplest conditions involve some specific affine functional Φ on V, related to $\mathscr{P}$ and $\mathscr{O}$, and the value that Φ takes at $x \in \mathscr{O}$.

Definition 1. A $\mathscr{P}^c$*-witness* is an affine functional Φ on V such that

(i) $\Phi(y) \leq 0$ for all $y \in \mathscr{P}$;
(ii) $\Phi(x) > 0$ for some $x \in \mathscr{P}^c$.

A $\mathscr{P}^c$-witness is *tight* if it satisfies the extra condition

(iii) $\Phi(z) = 0$ for some $z \in \mathscr{P}$.

If Φ is a $\mathscr{P}^c$-witness, the inequality $\Phi(x) > 0$ entails that $x \in \mathscr{O}$ does not possess the property $\mathscr{P}$. The set $W(\Phi) = \{x \in \mathscr{O} \mid \Phi(x) > 0\}$ is thus the subset of all objects in $\mathscr{P}^c$ *certified* by Φ. Note that there always exists a tight $\mathscr{P}^c$-witness

Φ' certifying at least as many objects in $\mathscr{P}^c$ as Φ. Namely, it is enough to set $\Phi'(v) = \Phi(v) - \max\{\Phi(y) \mid y \in \mathscr{P}\}$.

We now consider the following two natural questions:

- for any $v^* \in V^*$, find conditions ensuring the existence of $\mathscr{P}^c$-witnesses having linear part v^*;
- for a given $\mathscr{P}^c$-witness Φ, characterize all witnesses Φ' certifying the same states as Φ.

The first question has the following simple answer.

Proposition 1. *Let $v^* \in V^*$. There exist $\mathscr{P}^c$-witnesses having linear part v^* if and only if $\delta_{\mathscr{P}}(v^*) < \delta_{\mathscr{O}}(v^*)$. In this case, the affine functional $\Phi[v^*, \delta]$ is a $\mathscr{P}^c$-witness if and only if $\delta_{\mathscr{P}}(v^*) \leq \delta < \delta_{\mathscr{O}}(v^*)$, the equality being attained if and only if $\Phi[v^*, \delta]$ is tight.*

Proof. The affine functional $\Phi = \Phi[v^*, \delta]$ is a $\mathscr{P}^c$-witness if and only if, for all $y \in \mathscr{P}$ and some $x \in \mathscr{O} \setminus \mathscr{P}$,

$$\langle y , v^* \rangle - \delta \leq 0 < \langle x , v^* \rangle - \delta \quad \Leftrightarrow \quad \langle y , v^* \rangle \leq \delta < \langle x , v^* \rangle .$$

This is equivalent to $\delta_{\mathscr{P}}(v^*) \leq \delta < \delta_{\mathscr{O}}(v^*)$, where the equality is attained if and only if $\langle z , v^* \rangle = \delta$ for some $z \in \mathscr{P}$, that is, $\Phi(z) = 0$. □

Geometrically, the nonnegative gap $\delta_{\mathscr{O}}(v^*) - \delta_{\mathscr{P}}(v^*)$ is a measure of the distance between one of the two affine hyperplanes of V determined by v^* and touching $\mathscr{O}$ at its boundary, and the analogue hyperplane touching the boundary of $\mathscr{P}$. FIGURE?

The following quite natural fact gives a further insight of tightness.

Proposition 2. *If Φ is a tight $\mathscr{P}^c$-witness, then $\Phi(z_0) = 0$ for some extreme point z_0 of $\mathscr{P}$.*

Proof. The set $Z = \{z \in \mathscr{P} \mid \Phi(z) = 0\}$ is a nonempty closed and convex subset of $\mathscr{P}$. By Krein-Milman theorem [1, Cor. 18.5.1], Z has some extreme point z_0. We claim that z_0 is extreme also for $\mathscr{P}$. Indeed, suppose that $z_0 = \lambda y_1 + (1 - \lambda) y_2$ with $y_1, y_2 \in \mathscr{P}$ and $\lambda \in (0, 1)$. The conditions $0 = \Phi(z_0) = \lambda \Phi(y_1) + (1 - \lambda)\Phi(y_2)$ and $\Phi(y_i) \leq 0$ for $i = 1, 2$ then imply that $y_1, y_2 \in Z$, hence $y_1 = y_2 = z$. □

Now we come to the second problem. We say that two $\mathscr{P}^c$-witnesses Φ_1 and Φ_2 are *equivalent* if they certify the same elements of $\mathscr{P}^c$, that is, $W(\Phi_1) = W(\Phi_2)$. In this case, we write $\Phi_1 \approx \Phi_2$. If the subset $\mathscr{P}$ is sufficiently large in $\mathscr{O}$, we have the following characterization of equivalent $\mathscr{P}^c$-witnesses.

Proposition 3. *Suppose $\mathscr{P} \cap \mathrm{ri}(\mathscr{O}) \neq \emptyset$. Then, two $\mathscr{P}^c$-witnesses $\Phi[v_1^*, \delta_1]$ and $\Phi[v_2^*, \delta_2]$ are equivalent if and only if both conditions $v_1^* - \alpha v_2^* \in V(\mathscr{O})^\perp$ and $\delta_1 - \alpha\delta_2 = \langle v_0 , v_1^* - \alpha v_2^* \rangle$ hold for some $\alpha > 0$ and for some (hence for all) $v_0 \in \mathscr{M}(\mathscr{O})$.*

Proof. Denote $\Phi_i = \Phi[v_i^*, \delta_i]$. Moreover, fix any $x \in W(\Phi_1) = W(\Phi_1)$.
Since

$$\text{(o)} \qquad (\Phi_1 - \alpha\Phi_2)(v_0 + v) = \langle v , v_1^* - \alpha v_2^* \rangle + [\langle v_0 , v_1^* - \alpha v_2^* \rangle - (\delta_1 - \alpha\delta_2)]$$

we see that under the two conditions in the statement we have $\Phi_1(w) = \alpha\Phi_2(w)$ for all $w \in \mathscr{M}(\mathscr{O}) = v_0 + V(\mathscr{O})$. Hence, $\Phi_1 \approx \Phi_2$.
Conversely, assume $\Phi_1 \approx \Phi_2$. We preliminarly prove that the two nonempty sets $Z_i = \{z \in \mathscr{O} \mid \Phi_i(z) = 0\}$ $(i = 1, 2)$ actually coincide. Indeed, suppose by contradiction that $Z_1 \neq Z_2$. We can assume with no restriction that there is some $z \in \mathscr{O}$

such that $\Phi_1(z) = 0$ and $\Phi_2(z) \neq 0$, hence $\Phi_2(z) < 0$ since $W(\Phi_1) = W(\Phi_2)$. Then, $\Phi_1(\lambda x + (1-\lambda)z) = \lambda\Phi_1(x) > 0$, or, equivalently, $\Phi_2(\lambda x + (1-\lambda)z) > 0$ for all $\lambda \in (0,1)$, that contradicts continuity of the mapping $\lambda \mapsto \Phi_2(\lambda x + (1-\lambda)z)$.

We next claim that there is $z_0 \in \mathrm{ri}(\mathscr{O})$ such that $\Phi_i(z_0) = 0$ for all $i = 1,2$. To this aim, pick $y \in \mathscr{P} \cap \mathrm{ri}(\mathscr{O})$. Then, for all $i = 1,2$, either $\Phi_i(y) = 0$ and we are done, or $\Phi_i(y) < 0$. In the latter case, again by a continuity argument, there is some $\lambda \in (0,1)$ such that $\Phi_1(\lambda x + (1-\lambda)y) = \lambda\Phi_1(x) + (1-\lambda)\Phi_1(y) = 0$. Setting $z_0 = \lambda x + (1-\lambda)y$, we thus see that $z_0 \in Z_1 = Z_2$, and $z_0 \in \mathrm{ri}(\mathscr{O})$ by [1, Thm. 6.1].

By the usual rank-nullity theorem, z_0 can be completed to an affine basis $\{z_0, z_1, \ldots, z_m\}$ of $\mathscr{M}(\mathscr{O})$ such that $\Phi_1(z_k) = 0$ for all $k \leq m-1$ and $\Phi_1(z_m) > 0$. By possibly replacing the z_k's with $\mu_k z_k + (1-\mu_k)z_0$ for some $\mu_k \in (0,1)$, we can assume that $z_k \in \mathrm{ri}(\mathscr{O})$ for all $k = 1, \ldots, n$. Hence, also $\Phi_2(z_k) = 0$ for all $k \leq m-1$ and $\Phi_2(z_m) > 0$. It follows that

$$\left[\Phi_1 - \frac{\Phi_1(z_m)}{\Phi_2(z_m)}\Phi_2\right](z_0 + v) = 0 \qquad \forall z_0 + v \in z_0 + V(\mathscr{O}) = \mathscr{M}(\mathscr{O}),$$

and then (o) implies that $v_1^* - \alpha v_2^* \in V(\mathscr{O})^\perp$ and $\delta_1 - \alpha\delta_2 = \langle z_0 , v_1^* - \alpha v_2^* \rangle$ for $\alpha = \Phi_1(z_m)/\Phi_2(z_m) > 0$. $\square$

Corollary 1. *Suppose $\mathscr{P} \cap \mathrm{ri}(\mathscr{O}) \neq \emptyset$ and $\delta_{\mathscr{P}}(v_i^*) < \delta_{\mathscr{O}}(v_i^*)$ for $i = 1,2$. Then, the two tight $\mathscr{P}^c$-witnesses with respective linear parts v_1^* and v_2^* are equivalent if and only if $v_1^* - \alpha v_2^* \in V(\mathscr{O})^\perp$ for some $\alpha > 0$.*

Proof. Since $V(\mathscr{O})^\perp \subseteq V(\mathscr{P})^\perp$, if $v_1^* - \alpha v_2^* \in V(\mathscr{O})^\perp$, the condition $\delta_{\mathscr{P}}(v_1^*) - \alpha\delta_{\mathscr{P}}(v_2^*) = \langle v_0 , v_1^* - \alpha v_2^* \rangle$ for all $v_0 \in \mathscr{M}(\mathscr{O})$ automatically holds for the support function $\delta_{\mathscr{P}}$. $\square$

3. Incompatibility witnesses

We fix a finite dimensional complex Hilbert space $\mathcal{H}$, with $\dim \mathcal{H} = d$. We denote by $\mathcal{L}_{\mathrm{s}}(\mathcal{H})$ the real linear space of all selfadjoint operators on $\mathcal{H}$, endowed with the uniform operator norm $\|\cdot\|$. We write $\mathbb{1}$ for the identity operator. If Z is a finite set, we let $|Z|$ be its cardinality. An *observable* on a finite set Z is any map $\mathsf{C} : Z \to \mathcal{L}_{\mathrm{s}}(\mathcal{H})$ such that $\mathsf{C}(z) \geq 0$ for all $z \in Z$ and $\sum_{z\in Z} \mathsf{C}(z) = \mathbb{1}$. The *uniform observable* on Z is the observable U_Z with $\mathsf{U}_Z(z) = (1/|Z|)\,\mathbb{1}$ for all $z \in Z$.

All observables on Z constitute the closed and bounded convex subset $\mathrm{Obs}(Z)$ of the real linear space of all operator valued functions $H : Z \to \mathcal{L}_{\mathrm{s}}(\mathcal{H})$. We denote by $\mathrm{Fun}(Z)$ the latter linear space, and we assume it is endowed with the sup-norm $\|H\|_\infty = \max\{\|H(z)\| \mid z \in Z\}$; the dimension of $\mathrm{Fun}(Z)$ is $d^2\,|Z|$. If for any $A \in \mathcal{L}_{\mathrm{s}}(\mathcal{H})$ we define the affine set $\mathrm{Fun}_A(Z) = \{H \in \mathrm{Fun}(Z) \mid \sum_{Z\in Z} H(z) = A\}$, then $\mathrm{Obs}(Z) \subset \mathrm{Fun}_{\mathbb{1}}(Z)$.

Given two finite sets X and Y, two observables $\mathsf{A} \in \mathrm{Obs}(X)$ and $\mathsf{B} \in \mathrm{Obs}(Y)$ are *compatible* if there exists a third observable $\mathsf{C} \in \mathrm{Obs}(X \times Y)$ such that

$$\mathsf{C}_{[1]}(x) := \sum_{y\in Y} \mathsf{C}(x,y) \equiv \mathsf{A}(x) \qquad \text{and} \qquad \mathsf{C}_{[2]}(y) := \sum_{x\in X} \mathsf{C}(x,y) \equiv \mathsf{B}(y) \tag{2}$$

for all x and y. In this case, we say that (A,B) constitute a *compatible pair*. We denote by $\mathscr{O} := \mathrm{Obs}(X) \times \mathrm{Obs}(Y)$ the set of all pairs of observables on X and Y, and by $\mathscr{P}$ the subset of all compatible pairs in $\mathscr{O}$.

As it is well known, unless $\mathcal{H} = \mathbb{C}$ or $\min\{|X|, |Y|\} = 1$, the inclusion $\mathscr{P} \subset \mathscr{O}$ is strict. So, we will always assume this hypothesis in the following.

The sets $\mathscr{O}$ and $\mathscr{P}$ are convex and compact in the direct product linear space $V = \mathrm{Fun}(X) \times \mathrm{Fun}(Y)$; here, as the norm of V we choose the ℓ_∞-norm $\|(F,G)\|_\infty = \max\{\|F\|_\infty, \|G\|_\infty\}$. Indeed, only compactness of $\mathscr{P}$ needs to be checked; it follows by compactness of $\mathrm{Obs}(X \times Y)$ and continuity of the mapping $\mathrm{Obs}(X \times Y) \ni \mathsf{C} \mapsto (\mathsf{C}_{[1]}, \mathsf{C}_{[2]}) \in \mathrm{Obs}(X) \times \mathrm{Obs}(Y)$.

Definition 2. With the above definitions of $\mathscr{O}$, $\mathscr{P}$, V and $\mathscr{P}^c = \mathscr{O} \setminus \mathscr{P}$, an *incompatibility witness* is any $\mathscr{P}^c$-witness on the linear space V.

The next proposition gives some further insight into the convex structure of the sets $\mathscr{O}$ and $\mathscr{P}$.

Proposition 4. *The sets $\mathscr{O}$ and $\mathscr{P}$ defined above have the following properties.*
(a) $\mathscr{M}(\mathscr{P}) = \mathscr{M}(\mathscr{O}) = \mathrm{Fun}_{\mathbb{1}}(X) \times \mathrm{Fun}_{\mathbb{1}}(Y)$.
(b) $V(\mathscr{P}) = V(\mathscr{O}) = \mathrm{Fun}_0(X) \times \mathrm{Fun}_0(Y)$.
(c) $\mathrm{ri}(\mathscr{P}) \subset \mathrm{ri}(\mathscr{O})$ *and* $\mathsf{U}_{X\times Y} \in \mathrm{ri}(\mathscr{P})$.

Proof. Clearly,

$$(*) \quad \mathscr{M}(\mathscr{P}) \subseteq \mathscr{M}(\mathscr{O}) \subseteq \mathrm{Fun}_{\mathbb{1}}(X) \times \mathrm{Fun}_{\mathbb{1}}(Y) = (\mathsf{U}_X, \mathsf{U}_Y) + \mathrm{Fun}_0(X) \times \mathrm{Fun}_0(Y)\,.$$

In order to prove that the previous inclusions actually are equalities (item (a)), it is enough to show that

$$(**) \qquad (\mathsf{U}_X, \mathsf{U}_Y) + \varepsilon B_0 \subseteq \mathscr{P}$$

where

$$\varepsilon = \frac{1}{2} \min\left\{\frac{1}{|X|}, \frac{1}{|Y|}\right\}$$

and

$$B_0 = \{(F,G) \in \mathrm{Fun}_0(X) \times \mathrm{Fun}_0(Y) \mid \|(F,G)\|_\infty < 1\}$$

is the open unit ball in $\mathrm{Fun}_0(X) \times \mathrm{Fun}_0(Y)$. Indeed, if $(**)$ holds, then

$$(\mathsf{U}_X, \mathsf{U}_Y) + \mathrm{Fun}_0(X) \times \mathrm{Fun}_0(Y) = \mathscr{M}((\mathsf{U}_X, \mathsf{U}_Y) + \varepsilon B_0) \subseteq \mathscr{M}(\mathscr{P})\,.$$

Now, $(**)$ immediately follows since for $(F,G) \in B_0$, the formula

$$\mathsf{C}(x,y) = \mathsf{U}_{X\times Y} + \varepsilon \left[\frac{1}{|Y|} F(x) + \frac{1}{|X|} G(y)\right]$$

defines an element $\mathsf{C} \in \mathrm{Obs}(X \times Y)$ such that $(\mathsf{C}_{[1]}, \mathsf{C}_{[2]}) = (\mathsf{U}_X, \mathsf{U}_Y) + \varepsilon(F,G)$. Having proven that in $(*)$ all inclusions actually are equalities, item (b) is obvious, while item (c) follows from $(**)$ and [1, Cor. 6.5.2]. $\square$

The dual space V^* of $V = \mathrm{Fun}(X) \times \mathrm{Fun}(Y)$ can be identified with V itself by means of the pairing

$$\langle\, (F_1, G_1)\,,\, (F_2, G_2)\,\rangle = \sum_{x\in X} \mathrm{tr}\left[F_1(x) F_2(x)\right] + \sum_{y \in Y} \mathrm{tr}\left[G_1(y) G_2(y)\right]\,.$$

With this identification, formula (1) for the most general affine functional Φ on V can be rewritten $\Phi = \Phi[(F,G),\delta]$, with $(F,G) \in \mathrm{Fun}(X) \times \mathrm{Fun}(Y)$ and $\delta \in \mathbb{R}$.

The following class of incompatibility witnesses will play a very special role in in the next section.

Definition 3. An incompatibility witness Φ is *standard* if $\Phi = \Phi[(F,G),\delta]$, with
(i) $F(x) \geq 0$ and $G(y) \geq 0$ for all $x \in X$ and $y \in Y$;

(ii) $\sum_{x\in X} \operatorname{tr}[F(x)] + \sum_{y\in Y} \operatorname{tr}[G(y)] = 1$.

The next theorem shows that we can restrict to standard incompatibility witnesses with no loss of generality.

Theorem 1. *The following facts hold.*

(a) If $\Phi_1 = \Phi[(F_1, G_1), \delta_1]$ and $\Phi_2 = \Phi[(F_2, G_2), \delta_2]$ are two incompatibility witnessses, then $\Phi_1 \approx \Phi_2$ if and only if

$$F_2(x) = \alpha F_1(x) + A \qquad G_2(y) = \alpha F_2(y) + B \qquad \delta_2 = \alpha\delta_1 + \operatorname{tr}[A+B] \tag{3}$$

for some $\alpha > 0$ and $A, B \in \mathcal{L}_{\mathrm{s}}(\mathcal{H})$.

(b) For any incompatibility witness Φ, there exists a standard incompatibility witness Φ_{st} such that $\Phi \approx \Phi_{\mathrm{st}}$.

Proof. (a) Using the identification $V^* = V$ explained above, Prop. 4.(b) and a simple dimension counting lead to the equalities

$$\begin{aligned} V(\mathscr{P})^{\perp} = V(\mathscr{O})^{\perp} = \{(F,G) \in \operatorname{Fun}(X) \times \operatorname{Fun}(Y) \mid F(x_1) = F(x_2) \quad \text{and} \\ G(y_1) = G(y_2) \quad \forall x_1, x_2 \in X,\, y_1, y_2 \in Y\}\,. \end{aligned}$$

Moreover, Prop. 3 can be applied by Prop. 4.(c), hence the equivalence $\Phi_1 \approx \Phi_2$ holds if and only if, for some $\alpha > 0$, both the following conditions are satisfied:

- $(F_1, G_1) - \alpha(F_2, G_2) \in V(\mathscr{P})^{\perp} \;\Leftrightarrow\;$ there exist $A, B \in \mathcal{L}_{\mathrm{s}}(\mathcal{H})$ such that $F_1(x) = \alpha F_2(x) + A$ and $G_1(y) = \alpha G_2(y) + B$ for all $x \in X$, $y \in Y$;
- $\delta_1 - \alpha\delta_2 = \langle\, (\mathsf{U}_X, \mathsf{U}_Y)\,,\, (F_1, G_1) - \alpha(F_2, G_2)\,\rangle \;\Leftrightarrow\; \delta_1 = \alpha\delta_2 + \operatorname{tr}[A] + \operatorname{tr}[B]$ by the previous item.

This concludes the proof of (a).

(b) Let F_1, G_1 and δ_1 be such that $\Phi = \Phi[(F_1, G_1), \delta_1]$. Moreover, let μ be the minimal eigenvalue of all the operators $\{F_1(x) \mid x \in X\} \cup \{G_1(y) \mid y \in Y\}$. Then,

$$F_1(x) - \mu\mathbb{1} \geq 0 \quad \text{and} \quad G_1(y) - \mu\mathbb{1} \geq 0 \quad \text{for all } x \in X,\, y \in Y\,.$$

Since Φ is not constant on $\mathscr{O} \subset \mathscr{M}(\mathscr{O})$, it must be $(F_1, G_1) \notin V(\mathscr{O})^{\perp}$, hence $F_1(x) \neq \mu\mathbb{1}$ for some x or $G_1(y) \neq \mu\mathbb{1}$ for some y. It follows that

$$\frac{1}{\alpha} := \sum_{x\in X} \operatorname{tr}[F_1(x) - \mu\mathbb{1}] + \sum_{y\in Y} \operatorname{tr}[G_1(y) - \mu\mathbb{1}] > 0\,,$$

and the two operator valued functions

$$F_2(x) = \alpha(F_1(x) - \mu\mathbb{1}) \qquad\qquad G_2(y) = \alpha(G_1(y) - \mu\mathbb{1})$$

satisfy items (i) and (ii) of Definition 3. Set

$$\delta_2 = \alpha(\delta_1 - 2\mu d)\,.$$

Then, by item (a) above with $A = B = -\alpha\mu\mathbb{1}$, the affine functional $\Phi_{\mathrm{st}} = \Phi[(F_2, G_2), \delta_2]$ is a standard incompatibility witness that is equivalent to Φ. $\square$

4. A quick detour on conditioned quantum state discrimination

In the standard state discrimination scenario, Alice picks a label z, unknown to Bob, from a given set Z with probability $p(z)$, and she encodes it into a quantum state $\varrho_z \in \mathcal{L}_{\mathrm{s}}(\mathcal{H})$. She then sends the state ϱ_z to Bob, who performs a measurement of an observable $\mathsf{C} \in \mathrm{Obs}(Z)$ on it. Finally, Bob guesses the result of his measurement for the label z.

Both the probability distribution p on Z and Alice's encoding $z \mapsto \varrho_z$ are known to Bob. The objective is to choose the observable C in such a way to maximize Bob's probability of guessing the correct label:

$$P_{\mathrm{guess}} = \sum_{z\in Z} p(z)\mathrm{tr}\left[\varrho_z \mathsf{C}(z)\right] .$$

It is convenient to merge the probability distribution p and the encoding $z \mapsto \varrho_z$ into a single classical-quantum state, that is, an operator valued function $\mathcal{E} : Z \to \mathcal{L}_{\mathrm{s}}(\mathcal{H})$ such that

(i) $\mathcal{E}(z) \geq 0$ for all $z \in Z$;
(ii) $\sum_{z\in Z} \mathrm{tr}\left[\mathcal{E}(z)\right] = 1$.

We call such a function a *state ensemble* on Z. In this way, Bob's guessing probability can be rewritten

$$P_{\mathrm{guess}} \equiv P_{\mathrm{guess}}[\mathcal{E}](\mathsf{C}) = \sum_{z\in Z} \mathrm{tr}\left[\mathcal{E}(z)\mathsf{C}(z)\right] .$$

The aim of quantum state discrimination is to find the maximal guessing probability

$$P_{\mathrm{guess}}(\mathcal{E}) = \max_{\mathsf{C}\in\mathrm{Obs}(Z)} P_{\mathrm{guess}}[\mathcal{E}](\mathsf{C}) .$$

This scenario can be modified by allowing to transfer some classical information from Alice to Bob before of Bob's guessing. Namely, in quantum state discrimination with *pre-* or *post-measurement information*, Alice and Bob agree to partition the label set Z into two disjoint subsets X and Y, known to both of them. Everything then goes as above, with the fundamental difference that, either before or after Bob performs his measurement, Alice gives him the extra information about the subset X or Y she picked the state from. According to the time of Alice's announcement, the scenario splits into the following two situations.

4.1. State discrimination with post-measurement information.

Here, Alice announces the chosen subset X or Y *after* Bob has performed his measurement, but *before* he makes his guess. Thus, Bob is only allowed to post-process his measurement result according to Alice's extra information. No restriction is made on Bob's measured observable C; that is, he is free to choose the outcome set Z' of C as large as he wants, in order to store as much classical information as he needs for his post-processing. The set Z' may even be much larger than the original label set Z; indeed, it only matters that Z' is finally post-processed into either X or Y, in accordance with Alice's post-measurement information.

It turns out [2] that it is not restrictive to assume that $Z' = X \times Y$ and Bob's post-processing is the projection onto the factor declared by Alice. With this choice, the guessing probability obtained by post-processing an observable $\mathsf{C} \in \mathrm{Obs}(X \times Y)$

according to Alice's post-measurement information is

$$P_{\text{guess}}^{\text{post}}[\mathcal{E}](\mathsf{C}) = \sum_{x\in X} \operatorname{tr}\left[\mathcal{E}(x)\mathsf{C}_{[1]}(x)\right] + \sum_{y\in Y} \operatorname{tr}\left[\mathcal{E}(y)\mathsf{C}_{[2]}(y)\right] .$$

(see [2, Eq. (11)]). The maximal guessing probability in the post-measurement scenario is then

$$\begin{aligned} P_{\text{guess}}^{\text{post}}(\mathcal{E}) &= \max_{\mathsf{C}\in\mathrm{Obs}(X\times Y)} P_{\text{guess}}^{\text{post}}[\mathcal{E}](\mathsf{C}) \\ &= \max_{(\mathsf{A},\mathsf{B})\in\mathscr{P}} \left\{ \sum_{x\in X} \operatorname{tr}\left[\mathcal{E}(x)\mathsf{A}(x)\right] + \sum_{y\in Y} \operatorname{tr}\left[\mathcal{E}(y)\mathsf{B}(y)\right] \right\}, \end{aligned} \tag{4}$$

where $\mathscr{P}$ is the set of all compatible pairs of observables on X and Y defined in Sec. 3.

4.2. **State discrimination with pre-measurement information.** In this scenario, Alice's announces the subset she picked the state from *before* Bob performs his measurement. So, Bob's system is now described by one of the two conditioned state ensembles

$$\mathcal{E}_{[1]}(x) = \frac{1}{p(X)}\,\mathcal{E}(x) \qquad x\in X \qquad \text{or} \qquad \mathcal{E}_{[2]}(y) = \frac{1}{p(Y)}\,\mathcal{E}(y) \qquad y\in Y$$

according to Alice's pre-measurement information. Here,

$$p(X) = \sum_{x\in X} p(x) \qquad\qquad p(Y) = \sum_{y\in Y} p(y)\,,$$

and the *state subensembles* $\mathcal{E}_{[1]}$ and $\mathcal{E}_{[2]}$ are state ensembles on X and Y, respectively.

Then, based on Alice's announcement, Bob can decide wheter measuring an observable $\mathsf{A}\in\mathrm{Obs}(X)$ or $\mathsf{B}\in\mathrm{Obs}(Y)$, thus having success probability

$$P_{\text{guess}}^{\text{prior}}[\mathcal{E}](\mathsf{A},\mathsf{B}) = p(X)P_{\text{guess}}[\mathcal{E}_{[1]}](\mathsf{A}) + p(Y)P_{\text{guess}}[\mathcal{E}_{[2]}](\mathsf{B})\,.$$

In order to maximize it, Bob can optimize his measurement for the state subensemble at hand. Therefore, Bob's maximal guessing probability in the pre-measurement information scenario is

$$\begin{aligned} P_{\text{guess}}^{\text{prior}}(\mathcal{E}) &= p(X)P_{\text{guess}}(\mathcal{E}_{[1]}) + p(Y)P_{\text{guess}}(\mathcal{E}_{[2]}) \\ &= \max_{(\mathsf{A},\mathsf{B})\in\mathscr{O}} \left\{ \sum_{x\in X} \operatorname{tr}\left[\mathcal{E}(x)\mathsf{A}(x)\right] + \sum_{y\in Y} \operatorname{tr}\left[\mathcal{E}(y)\mathsf{B}(y)\right] \right\}, \end{aligned} \tag{5}$$

where $\mathscr{O}$ is the set of all pairs of observables on X and Y. Clearly,

$$P_{\text{guess}}^{\text{prior}}(\mathcal{E}) \geq P_{\text{guess}}^{\text{post}}(\mathcal{E})\,.$$

5. Conditioned quantum state discrimination and witnessing incompatibility

Given a state ensemble $\mathcal{E}$ on $Z = X \sqcup Y$, we can define a vector $(F_{\mathcal{E}}, G_{\mathcal{E}}) \in V = \mathrm{Fun}(X)\times\mathrm{Fun}(Y)$, given by

$$F_{\mathcal{E}}(x) = \mathcal{E}(x) \qquad G_{\mathcal{E}}(y) = \mathcal{E}(y) \qquad \forall x\in X,\, y\in Y\,.$$

Clearly, the components of $(F_{\mathcal{E}}, G_{\mathcal{E}})$ satisfy conditions (i) and (ii) of Definition 3. Thus, by Prop. 1, the affine functional

$$\Phi_{\mathcal{E}} = \Phi[(F_{\mathcal{E}}, G_{\mathcal{E}}), \delta_{\mathscr{P}}(F_{\mathcal{E}}, G_{\mathcal{E}})]$$

is a standard and tight incompatibility witness, provided that $\delta_{\mathscr{P}}(F_{\mathcal{E}}, G_{\mathcal{E}}) < \delta_{\mathscr{O}}(F_{\mathcal{E}}, G_{\mathcal{E}})$. Now, since

$$\langle\, (\mathsf{A}, \mathsf{B})\,,\, (F_{\mathcal{E}}, G_{\mathcal{E}})\,\rangle = \sum_{x\in X} \operatorname{tr}\left[\mathcal{E}(x)\mathsf{A}(x)\right] + \sum_{y\in Y} \operatorname{tr}\left[\mathcal{E}(y)\mathsf{B}(y)\right]\,,$$

the support functions $\delta_{\mathscr{O}}$ and $\delta_{\mathscr{P}}$ are nothing else than the optimal guessing probabilities (4) and (5) defined in Sec. 4. More precisely,

$$P^{\text{post}}_{\text{guess}}(\mathcal{E}) = \delta_{\mathscr{P}}(F_{\mathcal{E}}, G_{\mathcal{E}}) \qquad\qquad P^{\text{prior}}_{\text{guess}}(\mathcal{E}) = \delta_{\mathscr{O}}(F_{\mathcal{E}}, G_{\mathcal{E}})\,.$$

Thus, in a scenario where actually we do not know wheter Bob awaits Alice's information before performing the measurement, his guessing probability exceeding the bound $P^{\text{post}}_{\text{guess}}(\mathcal{E})$ is a witness of the fact that he performs his measurement *only after* Alice's announcement. In this case, in order to improve the post-measurement bound, Bob necessarily chooses his observable among two incompatible ones A and B. Equivalently, the inequality

$$\Phi_{\mathcal{E}}(\mathsf{A}, \mathsf{B}) > 0 \quad\Leftrightarrow\quad \langle\, (\mathsf{A}, \mathsf{B})\,,\, (F_{\mathcal{E}}, G_{\mathcal{E}})\,\rangle = P^{\text{prior}}_{\text{guess}}[\mathcal{E}](\mathsf{A}, \mathsf{B}) > P^{\text{post}}_{\text{guess}}(\mathcal{E})$$

witnesses the incompatibility of A and B.

As explained in Sec. 2, the gap $P^{\text{prior}}_{\text{guess}}(\mathcal{E}) - P^{\text{post}}_{\text{guess}}(\mathcal{E})$ can be regarded as a measure of the distance between two parallel affine hyperplanes of V, touching $\mathscr{O}$ and $\mathscr{P}$ at the respective boundaries.

Remark 1. It is not clear to us for which state ensembles $\mathcal{E}$ we have the strict inequality $P^{\text{prior}}_{\text{guess}}(\mathcal{E}) > P^{\text{post}}_{\text{guess}}(\mathcal{E})$. For example, if $\mathcal{E}_{[1]}(x) = p_1(x)|\varphi_x\rangle\langle\varphi_x|$ and $\mathcal{E}_{[2]}(y) = p_2(y)|\psi_y\rangle\langle\psi_y|$, where $\{\varphi_x \mid x \in X\}$ and $\{\psi_y \mid y \in Y\}$ are two orthonormal bases of $\mathcal{H}$, it is easy to check that $P^{\text{prior}}_{\text{guess}}(\mathcal{E}) > P^{\text{post}}_{\text{guess}}(\mathcal{E})$ if and only if $|\langle\, \varphi_x \,|\, \psi_y \,\rangle| < 1$ for some x, y. But is there some general necessary and sufficient condition?

References

[1] RT. Rockafellar, *Convex analysis.* Princeton University Press, Princeton, NJ (1997)
[2] C. Carmeli, T. Heinosaari, and A. Toigo, arXiv:1804.09693 (2018)