The theory of statistical comparison:

from majorization to the "quantum Blackwell theorem" and beyond

Francesco Buscemi (Nagoya University) Dipartimento di Matematica, Politecnico di Milano, 15 June 2023

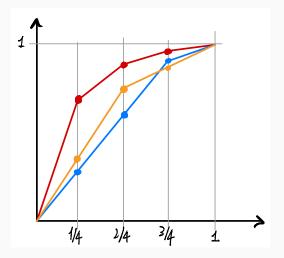
Prelude: majorization

Lorenz curves and majorization

- two probability distributions, $p = (p_1, \dots, p_n)$ and $q = (q_1, \dots, q_n)$
- truncated sums $P(k) = \sum_{i=1}^k p_i^{\downarrow}$ and $Q(k) = \sum_{i=1}^k q_i^{\downarrow}$, for all $k=1,\ldots,n$
- p majorizes q, i.e., $p >_{\text{maj}} q$, whenever $P(k) \geqslant Q(k)$, for all k
- minimal element: uniform distribution $e = n^{-1}(1, 1, \dots, 1)$

Hardy-Littlewood-Pólya, 1929

 $p >_{\text{maj}} q \iff q = Mp$, for some bistochastic matrix M.



$$(x_k, y_k) = (k/n, P(k)), \quad 1 \le k \le n$$

1/28

Blackwell's information preorder

Statistical experiments



Lucien Le Cam (1924-2000)

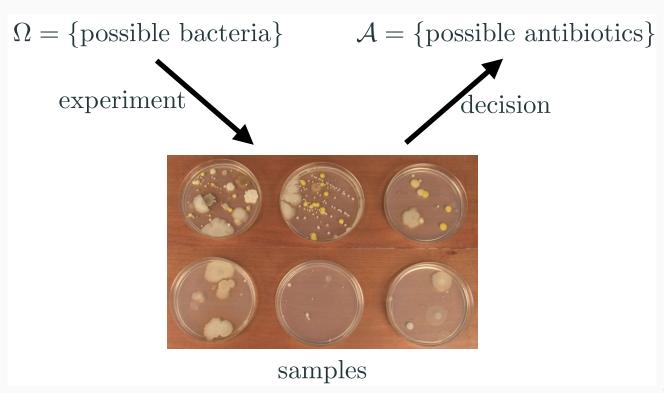
"The basic structures in the whole affair are systems that Blackwell called experiments, and transitions between them.

An experiment is a mathematical abstraction intended to describe the basic feature of an observational process if that process is contemplated in advance of its implementation."

Lucien Le Cam (1984)

2/28

A concrete example...



...and its abstract formulation

Definition (Statistical models and decision problems)

$$\Omega \xrightarrow{\text{experiment}} \mathcal{X} \xrightarrow{\text{decision}} \mathcal{A}$$

$$\begin{tabular}{ll} & & & \\ &$$

- parameter set $\Omega = \{\omega\}$, sample set $\mathcal{X} = \{x\}$, action set $\mathcal{A} = \{a\}$
- a statistical model/experiment is a triple $\mathbf{w} = \langle \Omega, \mathcal{X}, w(x|\omega) \rangle$
- a decision is a triple $\langle \mathcal{X}, \mathcal{A}, d(a|x) \rangle$
- a statistical decision problem/game is a triple $\mathbf{g} = \langle \Omega, \mathcal{A}, c \rangle$, where $c : \Omega \times \mathcal{A} \to \mathbb{R}$ is a payoff function

4/28

Playing statistical games with experiments

- the experiment/model is the resource: it is given
- the decision is the **transition**: it can be optimized

Ω	experiment	\mathcal{X}	decision	\mathcal{A}
\{		\{		\{
ω	$\overrightarrow{w(x \omega)}$	x	$\overrightarrow{d(a x)}$	a

Definition

The **(expected) maximin payoff** of a statistical model $\mathbf{w} = \langle \Omega, \mathcal{X}, w \rangle$ w.r.t. a decision problem $\mathbf{g} = \langle \Omega, \mathcal{A}, c \rangle$ is given by

$$c_{\mathbf{g}}^*(\mathbf{w}) \stackrel{\text{def}}{=} \max_{d(a|x)} \min_{\omega} \sum_{a,x} c(\omega,a) d(a|x) w(x|\omega)$$
.

Comparison of statistical models 1/2

For a fixed decision problem $\mathbf{g} = \langle \Omega, \mathcal{A}, c \rangle$, the payoffs $c_{\mathbf{g}}^*(\mathbf{w})$ and $c_{\mathbf{g}}^*(\mathbf{w}')$ can always be ordered (they are just real numbers).

6/28

Comparison of statistical models 2/2

Definition (Information preorder)

If the model $\mathbf{w} = \langle \Omega, \mathcal{X}, w \rangle$ is better than model $\mathbf{w}' = \langle \Omega, \mathcal{Y}, w' \rangle$ for all decision problems $\mathbf{g} = \langle \Omega, \mathcal{A}, c \rangle$, that is,

$$c_{\mathbf{g}}^*(\mathbf{w}) \geqslant c_{\mathbf{g}}^*(\mathbf{w}'), \quad \forall \mathbf{g} ,$$

then we say that \mathbf{w} is (always) more informative than \mathbf{w}' , and write

$$\mathbf{w} >_{\text{info}} \mathbf{w}'$$
.

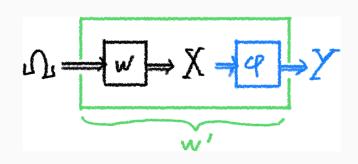
Can we visualize the information preorder more concretely?

Information preorder = statistical sufficiency

Theorem (Blackwell, 1953)

Given two statistical experiments $\mathbf{w} = \langle \Omega, \mathcal{X}, w \rangle$ and $\mathbf{w}' = \langle \Omega, \mathcal{Y}, w' \rangle$, the following are equivalent:

- 1. $\mathbf{w} >_{\text{info}} \mathbf{w}'$;
- 2. \exists cond. prob. dist. $\varphi(y|x)$ such that $w'(y|\omega) = \sum_{x} \varphi(y|x)w(x|\omega)$ for all y and ω .





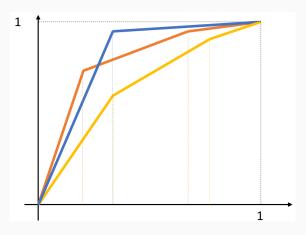
David Blackwell (1919-2010)

The case of dichotomies (a.k.a. relative majorization)

- for $\Omega=\{1,2\}$, we compare two dichotomies, i.e., two pairs of probability distributions $(\boldsymbol{p}_1,\boldsymbol{p}_2)$ and $(\boldsymbol{q}_1,\boldsymbol{q}_2)$, of dimension m and n, respectively
- \bullet relabel entries such that ratios p_1^i/p_2^i and q_1^j/q_2^j are nonincreasing
- for $\omega \in \{1,2\}$, let the truncated sums be $P_{\omega}(k) = \sum_{i=1}^k p_{\omega}^i$ and $Q_{\omega}(k) = \sum_{j=1}^k q_{\omega}^j$
- write $(p_1, p_2) >_{\text{maj}} (q_1, q_2)$ whenever the relative Lorenz curve of the former is never below that of the latter



For dichotomies, $\succ_{\mathrm{maj}} \iff \succ_{\mathrm{info}} \iff \exists$ stochastic matrix M s.t. $q_{\omega} = Mp_{\omega}$



$$(x_k, y_k) = (P_2(k), P_1(k)), \quad 1 \le k \le n$$

9/28

Quantum versions

Quantum statistical decision theory (Holevo, 1973)

classical case

- \bullet decision problems $\mathbf{g} = \langle \Omega, \mathcal{A}, c \rangle$
- models $\mathbf{w} = \langle \Omega, \mathcal{X}, \{w(x|\omega)\} \rangle$
- decisions d(a|x)
- $c_{\mathbf{g}}^{*}(\mathbf{w}) = \max_{d(a|x)} \min_{\omega} \cdots$

quantum case

- decision problems $\mathbf{g} = \langle \Omega, \mathcal{A}, c \rangle$
- quantum models $\mathcal{E} = \langle \Omega, \mathcal{H}_S, \{ \rho_S^{\omega} \} \rangle$
- POVMs $\{P_S^a: a \in \mathcal{A}\}$
- $c_{\mathbf{g}}^*(\mathcal{E}) = \max_{\{P_S^a\}} \min_{\omega} \sum_{a} c(\omega, a) \operatorname{Tr}[\rho_S^{\omega} P_S^a]$

10/28

Quantum statistical morphisms (FB, CMP 2012)

Definition (Generalized decisions)

Given a quantum statistical model (QSM) $\mathcal{E} = \langle \Omega, \mathcal{H}_S, \{\rho_S^\omega\} \rangle$, a family of operators $\{Z_S^a\}_a$ is said to be an \mathcal{E} -decision if and only if \exists POVM $\{P_S^a\}_a$ s.t.

$$\operatorname{Tr}[\rho_S^{\omega} Z_S^a] = \operatorname{Tr}[\rho_S^{\omega} P_S^a] , \quad \forall \omega, \forall a .$$

Definition (Statistical morphisms)

Given two QSMs $\mathcal{E} = \langle \Omega, \mathcal{H}_S, \{ \rho_S^{\omega} \} \rangle$ and $\mathcal{E}' = \langle \Omega, \mathcal{H}_{S'}, \{ \sigma_{S'}^{\omega} \} \rangle$, a linear map $\mathcal{M} : \mathsf{L}(\mathcal{H}_S) \to \mathsf{L}(\mathcal{H}_{S'})$ is said to be an $\mathcal{E} \to \mathcal{E}'$ quantum statistical morphism iff

- 1. \mathcal{M} is trace-preserving;
- 2. $\mathcal{M}(\rho_A^{\omega}) = \sigma_{S'}^{\omega}$, for all $\omega \in \Omega$;
- 3. the trace-dual map $\mathcal{M}^{\dagger}: \mathsf{L}(\mathcal{H}_{S'}) \to \mathsf{L}(\mathcal{H}_S)$ maps \mathcal{E}' -decisions into \mathcal{E} -decisions.

Quantum statistical comparison (FB, CMP 2012)

Given two QSMs $\mathcal{E} = \langle \Omega, \mathcal{H}_S, \{ \rho_S^{\omega} \} \rangle$ and $\mathcal{E}' = \langle \Omega, \mathcal{H}_{S'}, \{ \sigma_{S'}^{\omega} \} \rangle$

- information ordering: $\mathcal{E} >_{\inf} \mathcal{E}' \iff c_{\mathbf{g}}^*(\mathcal{E}) \geqslant c_{\mathbf{g}}^*(\mathcal{E}')$ for all \mathbf{g}
- complete information ordering: $\mathcal{E} \gg_{\inf} \mathcal{E}' \stackrel{\text{def}}{\iff} \mathcal{E} \otimes \mathcal{F} \succ_{\inf} \mathcal{E}' \otimes \mathcal{F}$ for all ancillary models $\mathcal{F} = \langle \Theta, \mathcal{H}_A, \{\tau_A^{\theta}\} \rangle$

Theorem 1/3: $\mathcal{E} >_{\mathrm{info}} \mathcal{E}'$ iff there exists a *quantum statistical morphism* $\mathcal{M} : \mathsf{L}(\mathcal{H}_S) \to \mathsf{L}(\mathcal{H}_{S'})$ such that $\mathcal{M}(\rho_S^\omega) = \sigma_{S'}^\omega$, $\forall \omega \in \Omega$

Theorem 2/3: $\mathcal{E} \gg_{\mathrm{info}} \mathcal{E}'$ iff there exists a *completely positive* trace-preserving linear map $\mathcal{N} : \mathsf{L}(\mathcal{H}_S) \to \mathsf{L}(\mathcal{H}_{S'})$ such that $\mathcal{N}(\rho_S^\omega) = \sigma_{S'}^\omega$ for all $\omega \in \Omega$

Theorem 3/3: if \mathcal{E}' is *commutative*, that is, if $[\sigma^{\omega_1}, \sigma^{\omega_2}] = 0$ for all $\omega_1, \omega_2 \in \Omega$, then $\mathcal{E} \gg_{\text{info}} \mathcal{E}'$ iff $\mathcal{E} >_{\text{info}} \mathcal{E}'$

12/28

Quantum dichotomies and quantum majorization

Classical hypothesis testing

- parameter set: $\Omega = \{1, 2\}$
- sample space: $\mathcal{X} = \{1, 2, \dots, n\}$
- ullet two possible hypotheses to test: $oldsymbol{p}_1$ or $oldsymbol{p}_2$
- an effect is a vector $\boldsymbol{t}=(t_1,t_2,\cdots,t_n)$ s.t. $0\leqslant t_i\leqslant 1$, $\forall i$
- a test is a pair $(\boldsymbol{t},\boldsymbol{e}-\boldsymbol{t})$, i.e., $\mathcal{A}\equiv\Omega$
- ullet expected probability of true positive: $oldsymbol{t} \cdot oldsymbol{p}_1$
- expected probability of false positive: $t \cdot p_2$

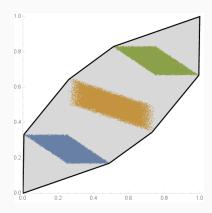
How to capture the "distinguishability" of the two hypotheses?

13/28

Relative testing region and relative Lorenz curve

The testing region of p_1 relative to p_2 is defined as the set

$$\mathcal{T}(\boldsymbol{p}_1\|\boldsymbol{p}_2) \stackrel{\text{def}}{=} \{(x,y) = (\boldsymbol{t}\cdot\boldsymbol{p_2},\boldsymbol{t}\cdot\boldsymbol{p_1}): \boldsymbol{t} \text{ effect}\}$$

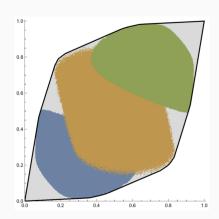


Theorem (Renes, JMP, 2016)

The relative Lorenz curve of $(\boldsymbol{p}_1,\boldsymbol{p}_2)$ coincides with the upper boundary of $\mathcal{T}(\boldsymbol{p}_1\|\boldsymbol{p}_2)$, so that $(\boldsymbol{p}_1,\boldsymbol{p}_2)>_{\mathrm{maj}}(\boldsymbol{q}_1,\boldsymbol{q}_2)\iff \mathcal{T}(\boldsymbol{p}_1\|\boldsymbol{p}_2)\supseteq\mathcal{T}(\boldsymbol{q}_1\|\boldsymbol{q}_2).$

14/28

Quantum relative Lorenz curve



Definition (FB and G. Gour, 2016)

Given two density matrices ρ_1 and ρ_2 on Hilbert space \mathcal{H} , the quantum testing region of ρ_1 relative to ρ_2 is defined as

$$\mathcal{T}(\rho_1 \| \rho_2) \stackrel{\text{def}}{=} \left\{ (x, y) = (\text{Tr}[E \ \rho_2], \text{Tr}[E \ \rho_1]) \right\},\,$$

where E can vary over all effects on $\mathcal H$ (i.e. $0 \le E \le 1$).

The quantum Lorenz curve of ρ_1 relative to ρ_2 is the upper boundary of $\mathcal{T}(\rho_1\|\rho_2)$ so that

$$(\rho_1, \rho_2) \succ_{\mathrm{maj}} (\sigma_1, \sigma_2) \stackrel{\mathsf{def}}{\iff} \mathcal{T}(\rho_1 \| \rho_2) \supseteq \mathcal{T}(\sigma_1 \| \sigma_2)$$

Remark. A quantum Lorenz curve may have strictly convex sections.

15/28

Equivalent characterizations of $>_{ m maj} 1/2$

Definition

Given two density matrices ρ and σ , we define the hypothesis testing relative entropy (FB, Datta; 2010)

$$D_H^{\epsilon}(\rho \| \sigma) := -\log \min_{\substack{0 \leqslant E \leqslant \mathbb{1} \\ \operatorname{Tr}[\rho E] \geqslant 1 - \epsilon}} \operatorname{Tr}[\sigma E] \ , \quad \epsilon \in [0, 1]$$

and the Hilbert α -divergence (FB, Gour; 2017)

$$H_{\alpha}(\rho \| \sigma) := \frac{\alpha}{\alpha - 1} \log \sup_{\frac{1}{\alpha} \mathbb{1} \leq E \leq \mathbb{1}} \frac{\operatorname{Tr}[\rho E]}{\operatorname{Tr}[\sigma E]}, \quad \alpha > 1,$$

with $H_1(\rho\|\sigma) := \lim_{\alpha \to 1^+} H_\alpha(\rho\|\sigma)$ and $H_\infty(\rho\|\sigma) := \lim_{\alpha \to \infty} H_\alpha(\rho\|\sigma)$

Equivalent characterizations of $>_{mai} 2/2$

Theorem (FB, Gour; 2017)

Given two quantum dichotomies (ρ, σ) and (ρ', σ') (possibly on different Hilbert spaces), the following are equivalent:

1.
$$\mathcal{T}(\rho \| \sigma) \supseteq \mathcal{T}(\rho' \| \sigma')$$
, i.e., $(\rho, \sigma) >_{\text{maj}} (\rho', \sigma')$

- 2. $D_H^{\epsilon}(\rho \| \sigma) \geqslant D_H^{\epsilon}(\rho' \| \sigma')$, for all $\epsilon \in [0, 1]$
- 3. $H_{\alpha}(\rho \| \sigma) \geqslant H_{\alpha}(\rho' \| \sigma')$ and $H_{\alpha}(\sigma \| \rho) \geqslant H_{\alpha}(\sigma' \| \rho')$, for all $\alpha \geqslant 1$
- 4. $\|\rho t\sigma\|_1 \ge \|\rho' t\sigma'\|_1$, for all $t \ge 0$

17/28

The problem with quantum dichotomies

classically:

$$(oldsymbol{p},oldsymbol{q})>_{ ext{maj}}(oldsymbol{p}',oldsymbol{q}')\stackrel{ ext{ ext{ ext{\sigma}}}}{\Longrightarrow}(oldsymbol{p},oldsymbol{q})>_{ ext{info}}(oldsymbol{p}',oldsymbol{q}')\stackrel{ ext{ ext{ ext{\sigma}}}}{\Longrightarrow}(oldsymbol{p},oldsymbol{q})\gg_{ ext{info}}(oldsymbol{p}',oldsymbol{q}')$$

the same equivalences hold also if both (ρ, σ) and (ρ', σ') are qubit dichotomies (Alberti and Uhlmann, 1980)

however, in general:

$$(\rho, \sigma) >_{\text{maj}} (\rho', \sigma') \stackrel{\longleftarrow}{\Rightarrow} (\rho, \sigma) >_{\text{info}} (\rho', \sigma') \stackrel{\longleftarrow}{\Rightarrow} (\rho, \sigma) \gg_{\text{info}} (\rho', \sigma')$$
 (counterexample by Matsumoto, 2014)

Problem: can we find conditions that are weaker, but easier to work with?

Information-theoretic treatment

Theorem (Matsumoto, 2010; FB, D. Sutter, M. Tomamichel, 2019)

Given two dichotomies (ρ, σ) and (ρ', σ') , if

$$D(\rho \| \sigma) > D(\rho' \| \sigma')$$
,

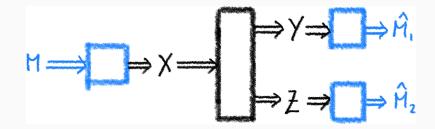
then there exists $\gamma > 0$, $n_0 \in \mathbb{N}$, and a sequence of CPTP linear maps $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$, such that

$$\begin{cases} \mathcal{E}_n(\sigma^{\otimes n}) = \sigma'^{\otimes n} & \forall n \in \mathbb{N} ,\\ \|\mathcal{E}_n(\rho^{\otimes n}) - \rho'^{\otimes n}\|_1 \leqslant e^{-\gamma n} & \forall n \geqslant n_0 . \end{cases}$$

19/28

Applications in information theory

Classical broadcast channels



How to capture the idea that Y carries more information than Z?

- (i) (stochastically) degradable: \exists channel $Y \rightarrow Z$
- (ii) less noisy: for all M, $H(M|Y) \leq H(M|Z)$
- (iii) less ambiguous: for all M, $\max \mathbb{P}\{\hat{M}_1 = M\} \geqslant \max \mathbb{P}\{\hat{M}_2 = M\}$
- (iv) less ambiguous (reformulation): for all M, $H_{\min}(M|Y) \leqslant H_{\min}(M|Z)$

Theorem (Körner-Marton, 1977; FB, 2016)

less noisy

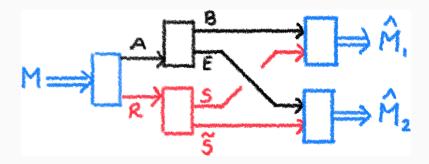
degradable

less ambiguous

ess ambiguous

20/28

Quantum broadcast channels



- (i) (CPTP) degradable: \exists channel $B \rightarrow E$
- (ii) completely less noisy: for all M and all symmetric side-channels $R\to S\tilde{S}$, $H(M|BS)\leqslant H(M|E\tilde{S})$
- (iii) completely less ambiguous: for all M and all symmetric side-channels $R \to S\tilde{S}, \ H_{\min}(M|BS) \leqslant H_{\min}(M|E\tilde{S})$

Theorem (FB-Datta-Strelchuk, 2014)

 $completely\ less\ noisy\ \stackrel{\Longrightarrow}{\ \Longleftrightarrow\ }\ degradable\ \Longleftrightarrow\ completely\ less\ ambiguous$

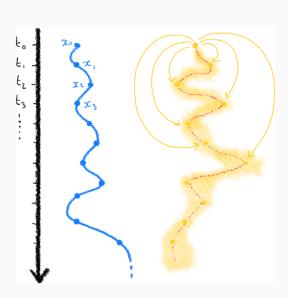
21/28

Applications in open quantum systems dynamics

Discrete-time stochastic processes

Formulation of the problem:

- for $i \in \mathbb{N}$, let x_i index the state of a system at time $t = t_i$
- given the system's initial state at time $t=t_0$, the process is fully predicted by the conditional distribution $p(x_N,\ldots,x_1|x_0)$
- if the system evolving is quantum, we only have a quantum dynamical mapping $\left\{\mathcal{N}_{Q_0 \to Q_i}^{(i)}\right\}_{i\geqslant 1}$
- the process is divisible if there exist channels $\mathcal{D}^{(i)}$ such that $\mathcal{N}^{(i+1)} = \mathcal{D}^{(i)} \circ \mathcal{N}^{(i)}$ for all $i \geqslant 1$
- **problem**: to provide a *fully information-theoretic* characterization of divisibility

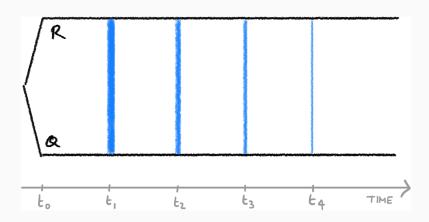


Divisibility as "information flow"

Theorem (FB-Datta, 2016; FB, 2018)

Given an initial open quantum system Q_0 , a quantum dynamical mapping $\left\{\mathcal{N}_{Q_0 \to Q_i}^{(i)}\right\}_{i \geqslant 1}$ is divisible if and only if, for any initial state ω_{RQ_0} ,

$$H_{\min}(R|Q_1) \leqslant H_{\min}(R|Q_2) \leqslant \cdots \leqslant H_{\min}(R|Q_N)$$
.



23/28

Applications in quantum thermodynamics

Quantum thermodynamics from relative majorization

Basic idea (FB, arXiv:1505.00535)

Thermal accessibility $\rho \to \sigma$ can be characterized as the statistical comparison between quantum dichotomies (ρ, γ) and (σ, γ) , for γ thermal state

Two main problems:

- for dimension larger than 2 and $[\sigma, \gamma] \neq 0$, we need a complete (i.e., extended) comparison
- moreover, Gibbs-preserving channels can create coherence between energy levels, while a truly thermal operation should not

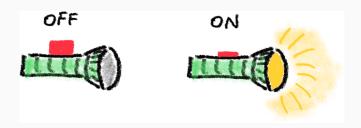
24/28

Complete comparison of quantum dichotomies 1/2

Definition (ON/OFF channels)

Given a d-dimensional quantum dichotomy $\mathcal{E}=(\rho,\gamma)$, we define the corresponding ON/OFF channel $\mathcal{N}_{\mathcal{E}}:\mathcal{L}(\mathbb{C}^2)\to\mathcal{L}(\mathbb{C}^d)$ as

$$\mathcal{N}_{\mathcal{E}}(\cdot) := \gamma \langle 0| \cdot |0\rangle + \rho \langle 1| \cdot |1\rangle$$

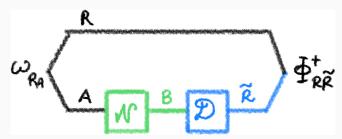


Complete comparison of quantum dichotomies 2/2

For a quantum channel $\mathcal{N}: A \to B$ and a state ω_{RA} , define the singlet fraction as

$$\Phi_{\omega}^*(\mathcal{N}) := \max_{\mathcal{D}: B \to \tilde{R}} \langle \Phi_{R\tilde{R}}^+ | (\mathsf{id}_R \otimes \mathcal{D} \circ \mathcal{N})(\omega_{RA}) | \Phi_{R\tilde{R}}^+ \rangle ,$$

where \mathcal{D} is a decoding quantum channel with output system $R \cong R$



Theorem (FB, 2015)

Given two quantum dichotomies $\mathcal{E}=(\rho_1,\rho_2)$ and $\mathcal{F}=(\sigma_1,\sigma_2)$, let $\mathcal{N}_{\mathcal{E}}$ and $\mathcal{N}_{\mathcal{F}}$ the corresponding ON/OFF channels. Then, $\mathcal{E} \gg \mathcal{F}$ if and only if

$$\Phi_{\omega}^*(\mathcal{N}_{\mathcal{E}}) \geqslant \Phi_{\omega}^*(\mathcal{N}_{\mathcal{F}}) , \quad \forall \omega$$

26/28

Dealing with quantum coherence (sketch)

For quantum dichotomies $\mathcal{E} = (\rho, \gamma)$ and $\mathcal{F} = (\sigma, \gamma)$ and group $\mathscr{T} = \{e^{-it\log\gamma}\}_{t\in\mathbb{R}}$, we write $\mathscr{E} \gg_{\mathscr{T}} \mathscr{F}$ iff \exists CPTP linear \mathscr{M} such that:

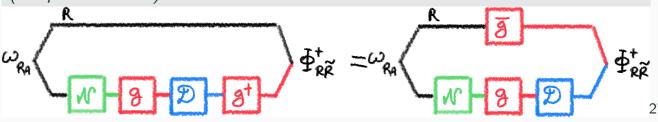
- (i) $\mathcal{M}(\rho) = \sigma$ and $\mathcal{M}(\gamma) = \gamma$;
- (ii) $\mathcal{M}(U_t \cdot U_t^{\dagger}) = U_t \mathcal{M}(\cdot) U_t^{\dagger}$, for all $t \in \mathbb{R}$

Theorem (Gour-Jennings-FB-Duan-Marvian, 2018)

 $\mathcal{E} \gg_{\mathscr{T}} \mathcal{F}$ if and only if

$$\widetilde{\Phi}_{\omega}^*(\mathcal{N}_{\mathcal{E}}) \geqslant \widetilde{\Phi}_{\omega}^*(\mathcal{N}_{\mathcal{F}}) , \quad \forall \omega$$

(see picture below)



Conclusions

Conclusions

- \bullet the theory of statistical comparison studies **morphisms** (preorders) of one "statistical object" X into another "statistical object" Y
- equivalent conditions are given in terms of (finitely or infinitely many) monotones, e.g., $f_i(X) \ge f_i(Y)$
- such monotones quantify the resources at stake in the operational framework at hand