From Statistical Decision Theory to Bell Nonlocality

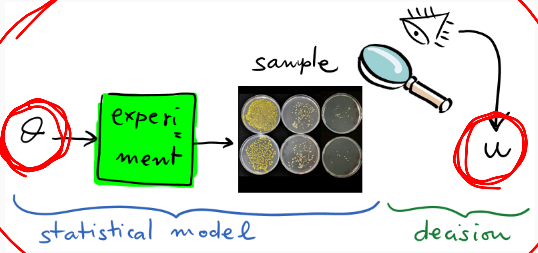
Francesco Buscemi*

QECDT, University of Bristol, 26 July 2018 (videoconference)

*Dept. of Mathematical Informatics, Nagoya University, buscemi@i.nagoya-u.ac.jp

Introduction

Statistical Decision Problems



$$\begin{array}{ccccc}
\Theta & \xrightarrow{\text{experiment}} & \mathcal{X} & \xrightarrow{\text{decision}} & \mathcal{U} \\
\rightsquigarrow & & \rightsquigarrow & & \rightsquigarrow \\
\theta & \xrightarrow[w(x|\theta)]{} & x & \xrightarrow[d(u|x)]{} & u \\
 & & & & \Downarrow \\
 & & & \mathcal{R} \ni & \ell(\theta, u)
\end{array}$$

Definition (Statistical Models and Decisions Problems)

A statistical experiment (i.e., statistical model) is a triple $\langle \Theta, \mathcal{X}, w \rangle$, a statistical decision problem (i.e., statistical game) is a triple $\langle \Theta, \mathcal{U}, \ell \rangle$.

How Much Is an Experiment Worth?

- the experiment *is given*, i.e., it is the "resource"
- the decision instead *can be optimized*

$$
\begin{tikzcd}
\Theta \arrow[r, "\text{experiment}"] \arrow[d, squiggly] & \mathcal{X} \arrow[r, "\text{decision}"] \arrow[d, squiggly] & \mathcal{U} \arrow[d, squiggly] \\
\theta \arrow[r, "w(x|\theta)"'] & x \arrow[r, "d(u|x)"'] & u \arrow[d, Rightarrow] \\
 & & \ell(\theta, u)
\end{tikzcd}
$$

Definition (Expected Payoff)

The expected payoff of a statistical experiment $\mathbf{w} = \langle \Theta, \mathcal{X}, w \rangle$ w.r.t. a decision problem $\langle \Theta, \mathcal{U}, \ell \rangle$ is given by

$$\mathbb{E}_{\langle \Theta, \mathcal{U}, \ell \rangle}[\mathbf{w}] \triangleq \max_{d(u|x)} \sum_{u,x,\theta} \ell(\theta, u) d(u|x) w(x|\theta) |\Theta|^{-1} .$$

Comparing Experiments 1/2

experiment $\mathbf{w} = \langle \Theta, \mathcal{X}, w(x|\theta) \rangle$

$$
\begin{tikzcd}
\Theta \arrow[r, "\text{experiment}"] \arrow[d, squiggly] & \mathcal{X} \arrow[r, "\text{decision}"] \arrow[d, squiggly] & \mathcal{U} \arrow[d, squiggly] \\
\theta \arrow[r, "w(x|\theta)"'] & x \arrow[r, "d(u|x)"'] & u \arrow[d, Rightarrow] \\
 & & \ell(\theta, u)
\end{tikzcd}
$$

experiment $\mathbf{w}' = \langle \Theta, \mathcal{Y}, w'(y|\theta) \rangle$

$$
\begin{tikzcd}
\Theta \arrow[r, "\text{experiment}"] \arrow[d, squiggly] & \mathcal{Y} \arrow[r, "\text{decision}"] \arrow[d, squiggly] & \mathcal{U} \arrow[d, squiggly] \\
\theta \arrow[r, "w'(y|\theta)"'] & y \arrow[r, "d'(u|y)"'] & u \arrow[d, Rightarrow] \\
 & & \ell(\theta, u)
\end{tikzcd}
$$

If $\mathbb{E}_{\langle \Theta, \mathcal{U}, \ell \rangle}[\mathbf{w}] \geq \mathbb{E}_{\langle \Theta, \mathcal{U}, \ell \rangle}[\mathbf{w}']$, then experiment $\langle \Theta, \mathcal{X}, w \rangle$ is better than experiment $\langle \Theta, \mathcal{Y}, w' \rangle$ for problem $\langle \Theta, \mathcal{U}, \ell \rangle$.

Comparing Experiments 2/2

Definition (Information Preorder)

If the experiment $\langle \Theta, \mathcal{X}, w \rangle$ is better than experiment $\langle \Theta, \mathcal{Y}, w' \rangle$ **for all** decision problems $\langle \Theta, \mathcal{U}, \ell \rangle$, then we say that $\langle \Theta, \mathcal{X}, w \rangle$ is *more informative* than $\langle \Theta, \mathcal{Y}, w' \rangle$, and write

$$\langle \Theta, \mathcal{X}, w \rangle \succeq \langle \Theta, \mathcal{Y}, w' \rangle \ .$$

Problem. The information preorder is operational, but not really "concrete". Can we visualize this better?

Blackwell's Theorem (1948-1953)

Blackwell-Sherman-Stein Theorem

Given two experiments with the same parameter space, $\langle \Theta, \mathcal{X}, w \rangle$ and $\langle \Theta, \mathcal{Y}, w' \rangle$, the condition $\langle \Theta, \mathcal{X}, w \rangle \succeq \langle \Theta, \mathcal{Y}, w' \rangle$ holds *iff* there exists a conditional probability $\varphi(y|x)$ such that $w'(y|\theta) = \sum_x \varphi(y|x) w(x|\theta)$.

$$\begin{tikzcd} \Theta \arrow[r] \arrow[d, squiggly] & \mathcal{Y} \arrow[d, squiggly] \\ \theta \arrow[r, "w'(y|\theta)"'] & y \end{tikzcd} = \begin{tikzcd} \Theta \arrow[r] \arrow[d, squiggly] & \mathcal{X} \arrow[r, "\text{noise}"] \arrow[d, squiggly] & \mathcal{Y} \arrow[d, squiggly] \\ \theta \arrow[r, "w(x|\theta)"'] & x \arrow[r, "\varphi(y|x)"'] & y \end{tikzcd}$$

David H. Blackwell (1919-2010)

An Important Special Case: Majorization

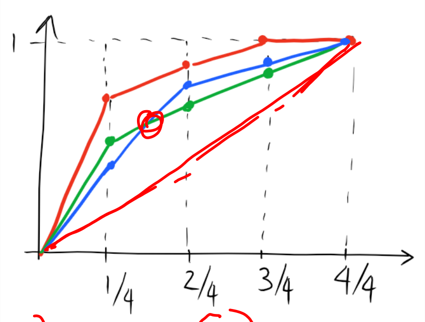
Lorenz Curves and Majorization Preorder

- two probability distributions, $\boldsymbol{p}$ and $\boldsymbol{q}$, of the same dimension n

$p^{\downarrow}$

- truncated sums $P(k) = \sum_{i=1}^{k} p_i^{\downarrow}$ and $Q(k) = \sum_{i=1}^{k} q_i^{\downarrow}$, for all $k = 1, \ldots, n$
- $\boldsymbol{p}$ majorizes $\boldsymbol{q}$, i.e., $\boldsymbol{p} \succeq \boldsymbol{q}$, whenever $P(k) \geq Q(k)$, for all k
- minimal element: uniform distribution $\boldsymbol{e} = n^{-1}(1, 1, \cdots, 1)$
- **Hardy, Littlewood, and Pólya (1929)**: $\boldsymbol{p} \succeq \boldsymbol{q} \iff \boldsymbol{q} = M\boldsymbol{p}$, for some bistochastic matrix M

Lorenz curve for probability distribution $\boldsymbol{p} = (p_1, \cdots, p_n)$:



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

Dichotomies and Tests

- a dichotomy is a statistical experiment with a two-point parameter space: $\langle \{1,2\}, \mathcal{X}, (\boldsymbol{w}_1, \boldsymbol{w}_2) \rangle$
- a testing problem (or "test") is a decision problem with a two-point action space $\mathcal{U} = \{1,2\}$

Definition (Testing Preorder)

Given two dichotomies $\langle \mathcal{X}, (\boldsymbol{w}_1, \boldsymbol{w}_2) \rangle$ and $\langle \mathcal{Y}, (\boldsymbol{w}'_1, \boldsymbol{w}'_2) \rangle$, we write

$$\langle \mathcal{X}, (\boldsymbol{w}_1, \boldsymbol{w}_2) \rangle \succeq_2 \langle \mathcal{Y}, (\boldsymbol{w}'_1, \boldsymbol{w}'_2) \rangle \quad ,$$

whenever

$$\mathbb{E}_{\langle \{1,2\}, \{1,2\}, \ell \rangle}[\langle \mathcal{X}, (\boldsymbol{w}_1, \boldsymbol{w}_2) \rangle] \geq \mathbb{E}_{\langle \{1,2\}, \{1,2\}, \ell \rangle}[\langle \mathcal{Y}, (\boldsymbol{w}'_1, \boldsymbol{w}'_2) \rangle]$$

for all testing problems.

Connection with Majorization Preorder

Blackwell's Theorem for Dichotomies (1953)

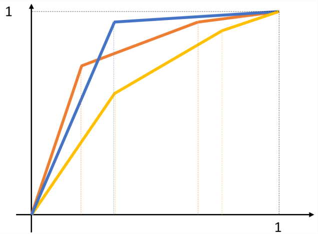
Given two dichotomies $\langle \mathcal{X}, (\boldsymbol{w}_1, \boldsymbol{w}_2) \rangle$ and $\langle \mathcal{Y}, (\boldsymbol{w}'_1, \boldsymbol{w}'_2) \rangle$, the relation $\langle \mathcal{X}, (\boldsymbol{w}_1, \boldsymbol{w}_2) \rangle \succeq_2 \langle \mathcal{Y}, (\boldsymbol{w}'_1, \boldsymbol{w}'_2) \rangle$ holds iff there exists a stochastic matrix M such that $M\boldsymbol{w}_i = \boldsymbol{w}'_i$.

- **majorization**: $\boldsymbol{p} \succeq \boldsymbol{q} \iff \langle \mathcal{X}, (\boldsymbol{p}, \boldsymbol{e}) \rangle \succeq_2 \langle \mathcal{X}, (\boldsymbol{q}, \boldsymbol{e}) \rangle$
- **thermomajorization**: as above, but replace uniform $\boldsymbol{e}$ with thermal distribution γ_T

Hence, the information preorder is a multivariate version of the majorization preorder, and Blackwell's theorem is a powerful generalization of that by Hardy, Littlewood, and Pólya.

Visualization: Relative Lorenz Curves

- two *pairs* of probability distributions, $(\boldsymbol{p}_1, \boldsymbol{p}_2)$ and $(\boldsymbol{q}_1, \boldsymbol{q}_2)$, of dimension m and n, respectively
- relabel their entries such that the ratios p_1^i / p_2^i and q_1^j / q_2^j are nonincreasing in i and j
- with such labeling, construct the truncated sums $P_{1,2}(k) = \sum_{i=1}^{k} p_{1,2}^i$ and $Q_{1,2}(k) = \sum_{j=1}^{k} q_{1,2}^i$
- $(\boldsymbol{p}_1, \boldsymbol{p}_2) \succeq_2 (\boldsymbol{q}_1, \boldsymbol{q}_2)$, if and only if the relative Lorenz curve of the former is never below that of the latter



Relative Lorenz curves:

$(x_k, y_k) = (P_2(k), P_1(k))$

The Quantum Case

Quantum Decision Theory (Holevo, 1973)

classical case	quantum case
• decision problems $\langle \Theta, \mathcal{U}, \ell \rangle$	• decision problems $\langle \Theta, \mathcal{U}, \ell \rangle$
• experiments $\mathbf{w} = \langle \Theta, \mathcal{X}, \{w(x\|\theta)\} \rangle$	• quantum experiments $\mathcal{E} = \langle \Theta, \mathcal{H}_S, \{\rho_S^\theta\} \rangle$
• decisions $d(u\|x)$	• POVMs $\{P_S^u : u \in \mathcal{U}\}$
• $p_c(u, \theta) = \sum_x d(u\|x) w(x\|\theta) \|\Theta\|^{-1}$	• $p_q(u, \theta) = \mathrm{Tr}\left[\rho_S^\theta \; P_S^u\right] \|\Theta\|^{-1}$
• $\mathbb{E}_{\langle \Theta, \mathcal{U}, \ell \rangle}[\mathbf{w}] = \max_{d(u\|x)} \sum \ell(\theta, u) p_c(u, \theta)$	• $\mathbb{E}_{\langle \Theta, \mathcal{U}, \ell \rangle}[\mathcal{E}] = \max_{\{P_S^u\}} \sum \ell(\theta, u) p_q(u, \theta)$

$< +\infty$

Hence, it is possible, for example, to compare quantum experiments with classical experiments, and introduce the information preorder as done before.

Example: Semiquantum Blackwell Theorem

Theorem (FB, 2012)

Given a quantum experiment $\mathcal{E} = \langle \Theta, \mathcal{H}_S, \{\rho_S^\theta\} \rangle$ and a classical experiment $\mathbf{w} = \langle \Theta, \mathcal{X}, \{w(x|\theta)\} \rangle$, the condition $\mathcal{E} \succeq \mathbf{w}$ holds *iff* there exists a POVM $\{P_S^x\}$ such that $w(x|\theta) = \mathrm{Tr}\left[P_S^x \; \rho_S^\theta\right]$.

Equivalent reformulation

Consider two quantum experiments $\mathcal{E} = \langle \Theta, \mathcal{H}_S, \{\rho_S^\theta\} \rangle$ and $\mathcal{E}' = \langle \Theta, \mathcal{H}_{S'}, \{\sigma_{S'}^\theta\} \rangle$, and assume that the σ's all commute. Then, $\mathcal{E} \succeq \mathcal{E}'$ holds *iff* there exists a quantum channel (CPTP map) $\Phi : \mathcal{L}(\mathcal{H}_S) \to \mathcal{L}(\mathcal{H}_{S'})$ such that $\Phi(\rho_S^\theta) = \sigma_{S'}^\theta$, for all $\theta \in \Theta$.

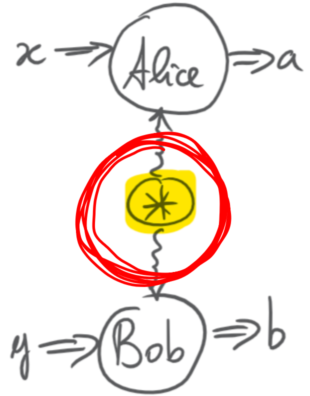
The Theory of Quantum Statistical Comparison

- fully quantum information preorder
- quantum relative majorization
- statistical comparison of quantum measurements (compatibility preorder)
- statistical comparison of quantum channels (input-degradability preorder, output-degradability preorder, simulability preorder, etc)
- applications: quantum information theory, quantum thermodynamics, open quantum systems dynamics, quantum resource theories, quantum foundations, . . .
- approximate case

Application to Quantum Foundations: Distributed Decision Problems, i.e., Nonlocal Games

Nonlocal Games

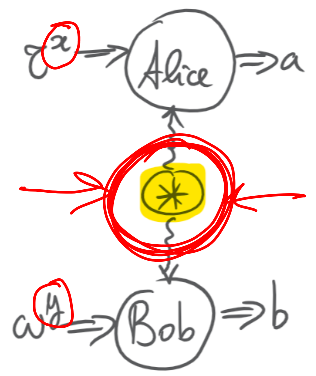
- nonlocal games (Bell tests) can be seen as bipartite decision problems $\langle \mathcal{X}, \mathcal{Y}; \mathcal{A}, \mathcal{B}; \ell \rangle$ played "in parallel" by non-communicating players
- with a classical source, $p_c(a, b|x, y) = \sum_\lambda \pi(\lambda) d_A(a|x, \lambda) d_B(b|y, \lambda)$
- with a quantum source, $p_q(a, b|x, y) = \mathrm{Tr}\left[\rho_{AB}(P_A^{a|x} \otimes Q_B^{b|y})\right]$



$$\mathbb{E}_{\langle \mathcal{X}, \mathcal{Y}; \mathcal{A}, \mathcal{B}; \ell \rangle}[*] \triangleq \max \sum_{x,y,a,b} \ell(x, y; a, b) p_{c/q}(a, b|x, y) |\mathcal{X}|^{-1} |\mathcal{Y}|^{-1}$$

Semiquantum Nonlocal Games

- semiquantum nonlocal games replace classical inputs with quantum inputs: $\langle \{\tau^x\}, \{\omega^y\}; \mathcal{A}, \mathcal{B}; \ell \rangle$
- with a classical source, $p_c(a,b|x,y) = \sum_\lambda \pi(\lambda) \operatorname{Tr}\left[(\tau_X^x \otimes \omega_Y^y) \left(P_X^{a|\lambda} \otimes Q_Y^{b|\lambda} \right) \right]$
- with a quantum source, $p_q(a,b|x,y) = \operatorname{Tr}\left[(\tau_X^x \otimes \rho_{AB} \otimes \omega_Y^y) \left(P_{XA}^a \otimes Q_{BY}^b \right) \right]$



$$\mathbb{E}_{\langle \{\tau^x\}, \{\omega^y\}; \mathcal{A}, \mathcal{B}; \ell \rangle}[*] \triangleq \max \sum_{x,y,a,b} \ell(x,y;a,b) p_{c/q}(a,b|x,y) |\mathcal{X}|^{-1} |\mathcal{Y}|^{-1}$$

Blackwell Theorem for Bipartite States

Theorem (FB, 2012)

Given two bipartite states ρ_{AB} and $\sigma_{A'B'}$, the condition (i.e., "nonlocality preorder")

$$\mathbb{E}_{\langle\{\tau^x\},\{\omega^y\};\mathcal{A},\mathcal{B};\ell\rangle}[\rho_{AB}] \geq \mathbb{E}_{\langle\{\tau^x\},\{\omega^y\};\mathcal{A},\mathcal{B};\ell\rangle}[\sigma_{A'B'}]$$

holds for all semiquantum nonlocal games, iff there exist CPTP maps $\Phi^{\lambda}_{A\to A'}$, $\Psi^{\lambda}_{B\to B'}$ and distribution $\pi(\lambda)$ such that

$$\sigma_{A'B'} = \sum_{\lambda} \pi(\lambda)\left(\Phi^{\lambda}_{A\to A'} \otimes \Psi^{\lambda}_{B\to B'}\right)(\rho_{AB}) .$$

Corollaries

- For any separable state ρ_{AB},

non

classical/sepe.

$$\mathbb{E}_{\langle\{\tau^x\},\{\omega^y\};\mathcal{A},\mathcal{B};\ell\rangle}[\rho_{AB}] = \mathbb{E}_{\langle\{\tau^x\},\{\omega^y\};\mathcal{A},\mathcal{B};\ell\rangle}[\rho_A \otimes \rho_B]$$
$$= \mathbb{E}^{\text{sep}}_{\langle\{\tau^x\},\{\omega^y\};\mathcal{A},\mathcal{B};\ell\rangle},$$

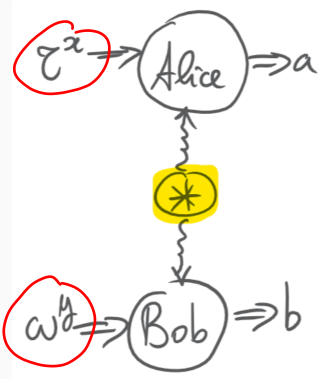
for all semiquantum nonlocal games.

- For any entangled state ρ_{AB}, **there exists** a semiquantum nonlocal game $\langle\{\tau^x\},\{\omega^y\};\mathcal{A},\mathcal{B};\ell\rangle$ such that

$$\mathbb{E}_{\langle\{\tau^x\},\{\omega^y\};\mathcal{A},\mathcal{B};\ell\rangle}[\rho_{AB}] > \mathbb{E}^{\text{sep}}_{\langle\{\tau^x\},\{\omega^y\};\mathcal{A},\mathcal{B};\ell\rangle}.$$

Other Properties of Semiquantum Nonlocal Games

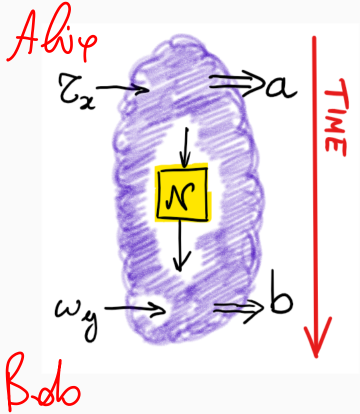
- can be considered as measurement device-independent entanglement witnesses (i.e., MDI-EW)
- can withstand losses in the detectors
- can withstand any amount of classical communication exchanged between Alice and Bob (not so conventional nonlocal games!)



Semiquantum Signaling Games

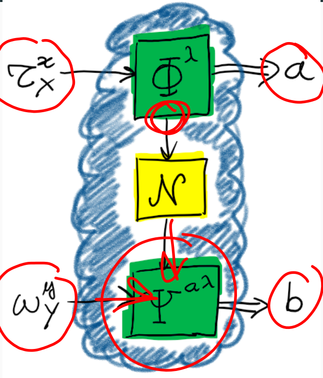
Semiquantum Nonlocality in Time

- turn dynamic communication into static memory!
- with unlimited classical memory, $p_c(a, b|x, y) = \sum_\lambda \pi(\lambda) \operatorname{Tr}\left[\tau_X^x P_X^{a|\lambda}\right] \operatorname{Tr}\left[\omega_Y^y Q_Y^{b|a,\lambda}\right]$
- if, moreover, a quantum memory $\mathcal{N} : A \to B$ is available, which correlations can be achieved?



Admissible Quantum Strategies

- τ_X^x is fed through an *instrument* $\{\Phi^{a|\lambda}_{X\to A}\}$, and outcome a is recorded
- the quantum output of the instrument is fed through the quantum memory $\mathcal{N} : A \to B$
- the output of the memory, together with ω_Y^y, are fed into a final measurement $\{\Psi^{b|a,\lambda}_{BY}\}$, and output b is recorded



$\sum_a \Phi^{a|\lambda}$

$$p_q(a,b|x,y) = \sum_\lambda \pi(\lambda)\,\mathrm{Tr}\left[\left(\{(\mathcal{N}_{A\to B}\circ\Phi^{a|\lambda}_{X\to A})(\tau_X^x)\}\otimes\omega_Y^y\right)\Psi^{b|a,\lambda}_{BY}\right]$$

Classical vs Quantum Strategies

Classical:

$$p_c(a,b|x,y) = \sum_\lambda \pi(\lambda)\,\mathrm{Tr}\left[\tau_X^x\; P_X^{a|\lambda}\right]\mathrm{Tr}\left[\omega_Y^y\; Q_Y^{b|a,\lambda})\right]$$

Quantum:

$$p_q(a,b|x,y) = \sum_\lambda \pi(\lambda)\,\mathrm{Tr}\left[\left(\{(\mathcal{N}_{A\to B}\circ\Phi_{X\to A}^{a|\lambda})(\tau_X^x)\}\otimes\omega_Y^y\right)\Psi_{BY}^{b|a,\lambda}\right]$$

Classical vs Quantum

Classical strategies correspond to the case in which the channel $\mathcal{N}$ has trivial output (completely depolarizing channel).

Statistical Comparison of Quantum Channels

Theorem (Rosset, FB, Liang, 2018)

Given two channels $\mathcal{N} : A \to B$ and $\mathcal{N}' : A' \to B'$, the condition (i.e., "signaling preorder")

$$\mathbb{E}_{\langle \{\tau^x\},\{\omega^y\};\mathcal{A},\mathcal{B};\ell\rangle}[\mathcal{N}] \geq \mathbb{E}_{\langle \{\tau^x\},\{\omega^y\};\mathcal{A},\mathcal{B};\ell\rangle}[\mathcal{N}']$$

holds for all semiquantum signaling games, iff there exist a quantum instrument $\{\Phi^a_{A'\to A}\}$ and CPTP maps $\Psi^a_{B\to B'}$ such that

$$\mathcal{N}'_{A'\to B'} = \sum_a \Psi^a_{B\to B'} \circ \mathcal{N}_{A\to B} \circ \Phi^a_{A'\to A} \; .$$

Consequences

- by asking quantum questions, it is possible to verify the quantumness in Alice's memory
- similar to Leggett-Garg inequalities, but without loopholes and other conceptual difficulties
- i.e., one of the simplest, non-trivial, time-like Bell tests

Conclusions

Conclusions

- generally speaking, the theory of statistical comparison studies transformation of one "statistical structure" X into another "statistical structure" Y
- equivalent conditions are given in terms of (finitely or infinitely many) *monotones*, e.g., $f_i(X) \geq f_i(Y)$
- such monotones shed light on the "resources" at stake in the operational framework at hand
- statistical comparison is complementary to SDP, which instead searches for *efficiently computable* functions like $f(X, Y)$
- however, SDP does not provide much insight into the resources at stake