# The Information-Disturbance Trade-Off in Quantum Theory

Francesco Buscemi (buscemi@i.nagoya-u.ac.jp)

Special Mathematics Lecture, 12 July 2018

# The Mechanical Certainty (Laplace's Demon)

*We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future, just like the past, would be present before its eyes.*

Pierre Simon Laplace, A Philosophical Essay on Probabilities (1814)

**Figure 1:** An *orrery* (clockwork reproducing the motion of planets).

**Quantum mechanics tells us that Laplace's dream is impossible not only in practice (complexity, chaos, etc)...**

**...but also in principle! Why?**

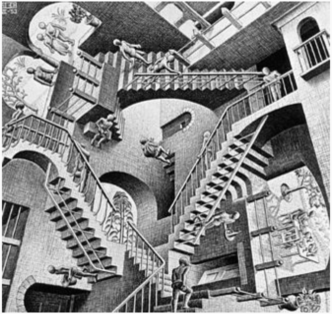
**The Heisenberg Principle, in its standard form, is not about the tradeoff between information and disturbance**

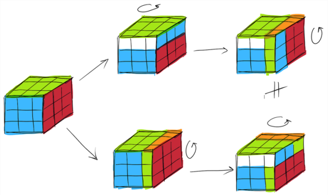
# Elementary Quantum Measurement Theory

- to each finite-state quantum system it is associated a finite-dimensional complex vector space $\mathbb{C}^d$ carrying an inner product $(\cdot, \cdot) : \mathbb{C}^d \times \mathbb{C}^d \to \mathbb{C}$
- pure states of the system are represented by unit-norm vectors $\psi \in \mathbb{C}^d$, i.e., $(\psi, \psi) = 1$
- measurable physical quantities (i.e., *observables*) are represented by self-adjoint operators $A = A^\dagger$, where $(\psi, A^\dagger \phi) \triangleq (A\psi, \phi)$
- a measurement of $A$ in state $\psi$ produces an outcome $a \in \sigma(A)$ with probability $\Pr\{A = a|\psi\} = (\psi, P_a \psi)$
- the average (i.e., expectation value) is defined as $\langle A \rangle_\psi \triangleq \sum_{a \in \sigma(A)} a \Pr\{A = a|\psi\} = (\psi, A\psi)$
- the variance is defined as $\mathrm{Var}_\psi(A) \triangleq \sqrt{\langle (A - \langle A \rangle_\psi)^2 \rangle_\psi} = \sqrt{\langle A^2 \rangle_\psi - \langle A \rangle_\psi^2}$
- given two observables $A$ and $B$, their commutator and anticommutator are defined as $[A, B] \triangleq AB - BA$ and $\{A, B\} \triangleq AB + BA$, respectively

# Non-commutative observables??

- operators in general do not commute: $AB \neq BA$
- it is easy to understand the meaning of noncommutativity for operators
- but what about the fact that two physical properties are noncommutative?
- in particular, what does noncommutativity mean in the context of quantum measurements? can we "witness" noncommutativity?





# The Heisenberg-Robertson Relation

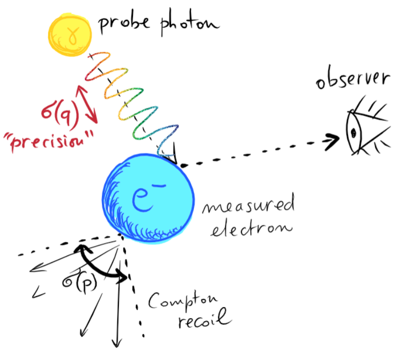
- assume that $(\psi, AB\psi) = z = a + ib$
- then, $2a = z + z^* = (\psi, AB\psi) + (AB\psi, \psi) = (\psi, AB\psi) + (\psi, (AB)^\dagger\psi) = (\psi, AB\psi) + (\psi, BA\psi) = \langle\{A, B\}\rangle_\psi$
- analogously, $2ib = z - z^* = \langle[A, B]\rangle_\psi$
- hence, $|z|^2 = a^2 + b^2 \geq b^2 = \frac{1}{4}|\langle[A, B]\rangle_\psi|^2 \geq 0$
- on the other hand, by the Cauchy-Schwarz inequality, $|z|^2 = |(\psi, AB\psi)|^2 = |(A\psi, B\psi)|^2 \leq \|A\psi\|^2\|B\psi\|^2$
- therefore, $\frac{1}{4}|\langle[A, B]\rangle_\psi|^2 \leq \|A\psi\|^2\|B\psi\|^2$
- by remapping $A \mapsto A - \langle A\rangle_\psi$ and $B \mapsto A - \langle B\rangle_\psi$, the commutator does not change, while $\|A\psi\|^2 \mapsto \mathrm{Var}_\psi^2(A)$

**Theorem (Heisenberg-Robertson relation)**

*For any two observables $A, B$ and any state $\psi$, we following bound holds:*

$$\mathrm{Var}_\psi(A)\mathrm{Var}_\psi(B) \geq \frac{1}{2}|\langle[A, B]\rangle_\psi| \, .$$

# The $\gamma$-Ray Microscope




Heisenberg in 1927 writes:

*Let $q_1$ be the precision with which the value $q$ is known (i.e., the mean error of $q$), therefore here the wavelength of the light. Let $p_1$ be the precision with which the value $p$ is determinable; that is, here, the discontinuous change of $p$ in the Compton effect (scattering). Then,*

$$p_1 q_1 \sim h \sim 10^{-34}\,\mathrm{Js}$$

Paraphrasing: the act of gathering information about the electron's position must cause an uncontrollable disturbance to the electron's momentum.

**Remark.** However, the Heisenberg-Robertson relation is not about the tradeoff between information and disturbance, but rather about the limitations in the preparation of quantum states.

# Let Us Begin with a Qualitative Statement...

## Basic Notions and Notations

In these slides:

- we label quantum systems by $Q, Q', \ldots$ and denote their (finite dimensional) Hilbert spaces $\mathcal{H}, \mathcal{H}', \ldots$
- the set of all linear operators on $\mathcal{H}$ is denoted $\mathsf{L}(\mathcal{H})$
- states are represented by density operators, i.e., $\rho \in \mathsf{L}(\mathcal{H})$ such that $\rho \succeq 0$ and $\mathrm{Tr}[\rho] = 1$
- we denote the set of all density operators on $\mathcal{H}$ as $\mathsf{D}(\mathcal{H})$
- linear maps from $\mathsf{L}(\mathcal{H})$ to $\mathsf{L}(\mathcal{H}')$ are denoted $\mathcal{E}, \mathcal{F}, \mathcal{R}, \ldots$; we usually assume that they are completely positive; the identity map is denoted id
- index sets (all finite) are denoted $\mathscr{A} = \{a\}$, $\mathscr{B} = \{b\}$, etc.
- classical random variables (usually thought as orthogonal states in a Hilbert space) are denoted $\mathbb{A}, \mathbb{X}$, etc.
- the maximally entangled state is denoted $|\tilde{\Phi}\rangle$
- we use the square fidelity $\mathsf{F}(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1^2$, which for pure states becomes $\mathsf{F}(|\psi\rangle, |\phi\rangle) = |\langle\psi|\phi\rangle|^2$

# What is a Measurement?

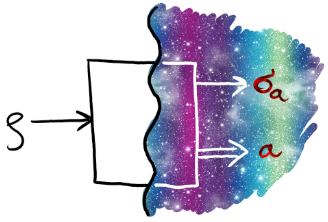
In operational theories, measurements are represented by families of linear maps, e.g., $\{\mathcal{E}_a : a \in \mathscr{A}\}$, indexed by the outcomes that can occur (index $a$). In quantum theory, there are some special requirements:

- for each $a$, the map $\mathcal{E}_a : \mathsf{L}(\mathcal{H}) \to \mathsf{L}(\mathcal{H}')$ is completely positive
- the sum $\sum_a \mathcal{E}_a$ is completely positive and trace-preserving

A family of operations like the one above is called **(completely positive) quantum instrument**.

### Operational Interpretation

Given that the state of the system immediately before the measurement is $\rho$, the outcome $a$ will be obtained with probability $p(a) \triangleq \mathrm{Tr}[\mathcal{E}_a(\rho)]$, in which case the state of the system immediately after the measurement will be $\sigma_a \triangleq \frac{1}{p(a)}\mathcal{E}_a(\rho)$.



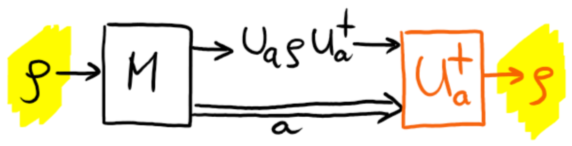
# Defining Disturbance (1/2)

**Definition (Naive Attempt)**

A measurement $\{\mathcal{E}_a\}_a$ is **non-disturbing** whenever, for any input $\rho$,

$$\mathcal{E}_a(\rho) \propto \rho, \qquad \forall a \in \mathscr{A} .$$

**Why this does not work.** Consider a measurement with $\mathcal{E}_a(\rho) = p(a) U_a \rho U_a^\dagger$. Even though $\mathcal{E}_a(\rho) \not\propto \rho$, knowing the outcome obtained, one can make this measurement non-disturbing by "undoing" the corresponding unitary transformation: $U_a^\dagger \mathcal{E}_a(\rho) U_a \propto \rho$.



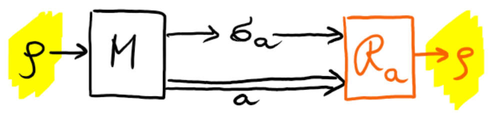
# Defining Disturbance (2/2)

The previous example tells us that disturbance is related to *irreversibility*, rather than state-change per se.

**Definition (Non-Disturbing Measurements)**

A measurement $\{\mathcal{E}_a\}_a$ is **physically non-disturbing** (viz., **physically reversible**) whenever there exists a family of CPTP linear maps $\{\mathcal{R}_a\}_a$ such that, for any input $\rho$,

$$(\mathcal{R}_a \circ \mathcal{E}_a)(\rho) \propto \rho, \qquad \forall a \in \mathscr{A} \ .$$




**Remark.** Notice the position of the universal quantifiers: the *same family* of correction operations $\{\mathcal{R}_a\}_a$ must be able to reverse the measurement process *for any possible* input state $\rho$.

**Remark.** Notice the difference between the measurement $\{\mathcal{E}_a\}_a$ and the correction $\{\mathcal{R}_a\}_a$: the former is a family of CP maps, which need not be TP, but whose sum is TP; the latter is a family of CPTP maps.

# Defining Information (or the Lack Thereof)

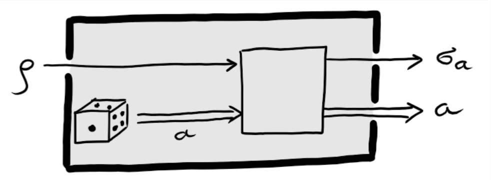
The information gained in a measurement resides in the way the outcomes are distributed.

**Definition (Uninformative Measurements)**

A measurement $\{\mathcal{E}_a\}_a$ is **uninformative** whenever the outcome probability distribution $p(a)$ does not depend on the input, in formula,

$$\mathrm{Tr}[\mathcal{E}_a(\rho)] = p(a), \qquad \forall \rho \ .$$

Hence, an uninformative measurement returns an outcome chosen at random, without even looking at the input state.




**Remark.** The output state could still depend on the input: the point is that the outcome $a$ does not!

# All Physically Reversible Measurements Are Uninformative

A simple consequence of the linearity of maps $\mathcal{E}_a$ and $\mathcal{R}_a$ is the following

**Theorem (No Information Without Disturbance, Part 1)**

*If a measurement $\{\mathcal{E}_a\}_a$ is physically reversible, then it is uninformative.*

**Proof.**

1. There exist CPTP $\{\mathcal{R}_a\}_a$ such that $(\mathcal{R}_a \circ \mathcal{E}_a)(\rho) \propto \rho$ for all $\rho$ and all $a$
2. Suppose that there exist two linearly independent states $\rho, \sigma$, such that $(\mathcal{R}_a \circ \mathcal{E}_a)(\rho) = p(a)\rho$ and $(\mathcal{R}_a \circ \mathcal{E}_a)(\sigma) = q(a)\sigma$, with $p(a) \neq q(a)$
3. Since $(\rho + \sigma)/2$ is also a state, point 1 implies $(\mathcal{R}_a \circ \mathcal{E}_a)(\rho + \sigma) = r(a)(\rho + \sigma)$
4. However, by linearity, we also have $(\mathcal{R}_a \circ \mathcal{E}_a)(\rho + \sigma) = p(a)\rho + q(a)\sigma$
5. Hence, $\{r(a) - p(a)\}\rho = \{q(a) - r(a)\}\sigma$
6. Since $\rho, \sigma$ are lin. indip., this implies $r(a) - p(a) = q(a) - r(a) = 0$, that is, $p(a) = q(a) = r(a)$
7. Contradiction with point 2

Hence, if the measurement is physically reversible, the proportionality coefficients $(\mathcal{R}_a \circ \mathcal{E}_a)\rho = p(a)\rho$ are the same for any $\rho$. Thus, since the maps $\mathcal{R}_a$ are all TP, the measurement is uninformative. $\square$

# Stochastic Reversibility

- In the previous proof, we only used linearity, never invoking complete positivity nor the Hilbert space structure. It is thus very general and it indeed holds for most operational theories, including classical probability theory!
- The reason is that physical reversibility is a very strong condition, as it must hold for each outcome. In quantum information theory one is often interested in an average (stochastic) condition.

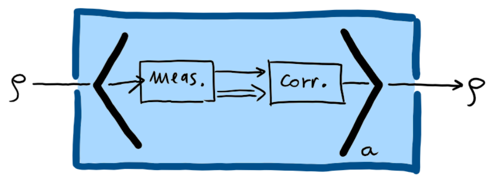
**Definition (Stochastically Reversible Measurements)**

A measurement $\{\mathcal{E}_a\}_a$ is **stochastically reversible** whenever there exists a family of CPTP linear maps $\{\mathcal{R}_a\}_a$ such that

$$\sum_{a\in\mathscr{A}}(\mathcal{R}_a\circ\mathcal{E}_a)(\rho)=\rho, \qquad \forall\rho\in\mathsf{D}(\mathcal{H})\ .$$

# Physical Reversibility vs Stochastic Reversibility

| Physical Reversibility | Stochastic Reversibility |
|---|---|
| there exist CPTP maps $\{\mathcal{R}_a\}_a$ such that $(\mathcal{R}_a \circ \mathcal{E}_a)(\rho) \propto \rho$ for all $a$ and all $\rho$ | there exist CPTP maps $\{\mathcal{R}_a\}_a$ such that $\sum_a (\mathcal{R}_a \circ \mathcal{E}_a)(\rho) = \rho$ for all $\rho$ |




Hence, any physically reversible measurement is also stochastically so, but not vice versa.

**Remark.** The terminology "physically reversible" vs "stochastically reversible" is taken from the analogous definition of "physically degradable" vs "stochastically degradable" for noisy channels in classical information theory.

# All Stochastically Reversible Measurements Are Uninformative

**Theorem (No Information Without Disturbance, Part 2)**

*In quantum theory, if a measurement $\{\mathcal{E}_a\}_a$ is stochastically reversible, then it is also physically reversible and, hence, uninformative.*

**Proof.**

1. The condition $\sum_a (\mathcal{R}_a \circ \mathcal{E}_a)(\rho) = \rho$, applied to a complete set of states, gives $\sum_a \mathcal{R}_a \circ \mathcal{E}_a = \mathrm{id}$
2. Hence, using the Choi-Jamiołkowski isomorphism between channels and bipartite states, $\left[\mathrm{id} \otimes \sum_a (\mathcal{R}_a \circ \mathcal{E}_a)\right] \left(|\tilde{\Phi}\rangle\langle\tilde{\Phi}|\right) = |\tilde{\Phi}\rangle\langle\tilde{\Phi}|$
3. Since $|\tilde{\Phi}\rangle\langle\tilde{\Phi}|$ is pure, it must be that $\left[\mathrm{id} \otimes (\mathcal{R}_a \circ \mathcal{E}_a)\right] \left(|\tilde{\Phi}\rangle\langle\tilde{\Phi}|\right) \propto |\tilde{\Phi}\rangle\langle\tilde{\Phi}|$, $\forall a$
4. Equivalently, $\mathcal{R}_a \circ \mathcal{E}_a \propto \mathrm{id}$, $\forall a$
5. Hence, the measurement $\{\mathcal{E}_a\}_a$ is physically reversible

□

**Remark.** Notice how here we made use of the full structure provided by quantum theory (e.g., complete positivity in point 2). Indeed, the above theorem does not hold in classical probability theory.

# Some Comments

- The above theorems only describe a qualitative tradeoff: measurements that are *exactly* reversible must be *exactly* uninformative
- Since in practice nothing is "exact," it is important to understand how information and disturbance are related in general
- For example, can we prove something like "If a measurement is *almost* reversible then it must be *almost* uninformative"? If yes, with respect to what measure is "almost" defined?

# Quantum Disturbance and Quantum Information Gain

# How to Quantify Reversibility

**Definition (Reversibility Index)**

Given a measurement $\mathcal{M} = \{\mathcal{E}_a\}_a$, we define its **(isotropic) reversibility index** as

$$R(\mathcal{M}) \triangleq \max \langle\tilde{\Phi}| \left\{ \left[ \mathrm{id} \otimes \sum_a (\mathcal{R}_a \circ \mathcal{E}_a) \right] \left( |\tilde{\Phi}\rangle\langle\tilde{\Phi}| \right) \right\} |\tilde{\Phi}\rangle ,$$

where the maximum is taken over all families of CPTP correction operations $\{\mathcal{R}_a\}_a$.

**Remark.** The reversibility index is equal to the (square) fidelity between the maximally entangled state and the Choi-Jamiołkowski state corresponding to $\sum_a (\mathcal{R}_a \circ \mathcal{E}_a)$. Thus, it is equal to one if and only if the measurement is *stochastically* reversible.

**Remark.** The reversibility index $R$, if high, guarantees that *any initial pure state* can be recovered, in average, with high accuracy: if $\{\overline{\mathcal{R}}_a\}_a$ are the operations achieving the maximum in the definition,

$$\int \mathrm{d}\,\psi \; \langle\psi| \sum_a (\overline{\mathcal{R}}_a \circ \mathcal{E}_a)(|\psi\rangle\langle\psi|)|\psi\rangle \geq R(\mathcal{M}) ,$$

where $\mathrm{d}\,\psi$ is the uniform (Haar invariant) measure over pure states.

# How to Quantify Information

- Information is always *about* something: for example, an arbitrarily chosen orthonormal basis (a "context") $\{|v_x\rangle\}_{x=1}^d$
- For such a choice, we compute the correlation (input/output joint distribution) $p(x,a) = d^{-1}\,\mathrm{Tr}[\mathcal{E}_a(|v_x\rangle\langle v_x|)]$
- Then, the mutual information $I(\mathbb{X};\mathbb{A}) = H(\mathbb{X}) + H(\mathbb{A}) - H(\mathbb{X}\mathbb{A})$ is a good measure of the average information that the outcome index $a$ contains about the input label $x$

However, in a quantum system, an infinite choice of bases is possible. Hence, we are led to the following

**Definition (Informational Power)**

Given a measurement $\mathcal{M} = \{\mathcal{E}_a\}_a$, we define its **informational power** as

$$I(\mathcal{M}) \triangleq \max I(\mathbb{X};\mathbb{A}) \ ,$$

where the maximum is taken over all choices of orthonormal bases* $\{|v_x\rangle\}_x$.

*: this is somehow a simplification; the maximization should run over all ensembles, not only orthonormal bases.

# Some Comments

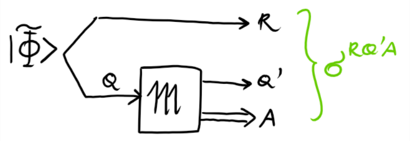
Two problems with the present formulation:

- While the informational power is an information-theoretic measure (defined in terms of Shannon entropies), the reversibility index is not (it's a fidelity)
- Both the informational power and the reversibility index involve a difficult optimization problem

We address both problems in what follows.

# Quantum Disturbance and Quantum Information Gain

Introducing a "reference" $R$, maximally entangled with $Q$, we have a tripartite configuration as follows:




for $\sigma^{RQ'\mathbb{A}} = \sum_a p(a)\sigma_a^{RQ'} \otimes |a\rangle\langle a|^{\mathbb{A}}$ and $p(a)\sigma_a^{RQ'} = (\mathrm{id}^R \otimes \mathcal{E}_a^Q)(|\tilde{\Phi}\rangle\langle\tilde{\Phi}|^{RQ})$

**Definition (Quantum Information Gain and Quantum Disturbance)**

Given a measurement $\mathcal{M} = \{\mathcal{E}_a\}_a$, we define its **quantum information gain** as

$$\iota(\mathcal{M}) \triangleq I(R;\mathbb{A}) = \log d - \sum_a p(a) S(\sigma_a^R) \ ,$$

and its **quantum disturbance** as

$$\delta(\mathcal{M}) \triangleq \log d - \underbrace{[S(\sigma^{Q'\mathbb{A}}) - S(\sigma^{RQ'\mathbb{A}})]}_{I_c^{R\to Q'\mathbb{A}}(\sigma^{RQ'\mathbb{A}})} \ .$$

# Why Such Names?

Why "quantum information gain"?

- Because

$$I(\mathcal{M}) \le \iota(\mathcal{M}) \le f_1(I(\mathcal{M}))\ , \qquad \text{where } \lim_{x\to 0} f_1(x) = 0$$

- Moreover, $\iota(\mathcal{M})$ is the optimal compression rate in Winter's measurement compression protocol, and it is closely related with Groenewold's information gain (1971)

Why "quantum disturbance"? Because [Schumacher and Westmoreland, QIP 2002; Junge et al, 2015]

$$-\log_2 R(\mathcal{M}) \le \delta(\mathcal{M}) \le f_2(1 - R(\mathcal{M}))\ , \qquad \text{where } \lim_{x\to 0} f_2(x) = 0$$

Hence, the quantum information gain $\iota(\mathcal{M})$ and the quantum disturbance $\delta(\mathcal{M})$ are equivalent to the informational power and the (ir)reversibility index, respectively; however, they do not involve any optimization and can be readily computed given the measurement $\{\mathcal{E}_a\}_a$.

# No (Large) Information Without (Large) Disturbance

**Theorem (Global Tradeoff)**

*For any measurement $\mathcal{M} = \{\mathcal{E}_a\}_a$, the information-disturbance tradeoff relation*

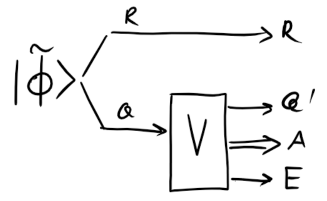
$$\delta(\mathcal{M}) \geq \iota(\mathcal{M})$$

*holds.*

**Proof.**

- Construct the "channelization" of the measurement $\mathcal{M}(\rho) \triangleq \sum_a \mathcal{E}_a(\rho) \otimes |a\rangle\langle a|^{\mathbb{A}}$
- Its Stinespring-Kraus dilation $V$ can be written as $|\tilde{\Phi}\rangle \to \sum_a |\Psi_a\rangle^{RQ'E_1}|a\rangle^{\mathbb{A}}|a\rangle^{E_2}$, where $E = E_1E_2$ is the environment
- Then, $\delta(\mathcal{M}) = S(R) - S(Q'\mathbb{A}) + S(RQ'\mathbb{A}) = S(R) - S(RE_1E_2) + S(E_1E_2) = I(R;E_1E_2) = I(R;E_1\mathbb{A}) \geq I(R;\mathbb{A}) = \iota(\mathcal{M})$

$\square$



# We can learn about the present, but at the cost of being unable to fully predict the future: Laplace's demon is defeated!