The Information-Disturbance Tradeoff in Quantum Theory

Francesco Buscemi[1]

Guest Lecture at the Department of Physics
National Cheng Kung University, Tainan
8 November 2017

[1]Dept. of Mathematical Informatics, Nagoya University, buscemi@i.nagoya-u.ac.jp
slides available at https://tinyurl.com/BTL20171108

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?

Let Us Begin with a Qualitative Statement...

Basic Notions and Notations

In these slides:

- we label quantum systems by $Q, Q', \dots$ and denote their (finite dimensional) Hilbert spaces $\mathcal{H}, \mathcal{H}', \dots$
- 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 $\mathcal{E}, \mathcal{F}, \mathcal{R}, \dots$; 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?

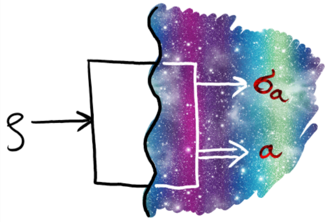
In operational theories, measurements are represented by families of operations, e.g., $\{\mathcal{E}_a : a \in \mathscr{A}\}$, indexed by the outcomes that can occur (index a). In quantum theory, there are some special requirements:

- 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 ρ, 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 $\sigma_a \triangleq \frac{1}{p(a)}\mathcal{E}_a(\rho)$.



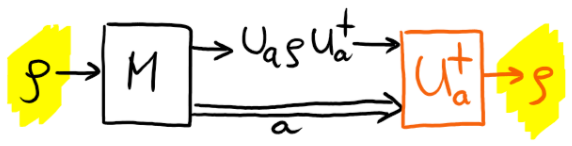
Defining Disturbance (1/2)

Definition (Naive Attempt)

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

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

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: $U_a^\dagger \mathcal{E}_a(\rho) U_a \propto \rho$.



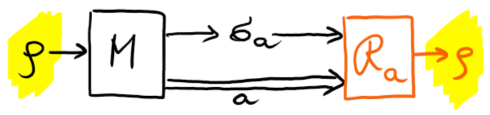
Defining Disturbance (2/2)

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 ρ,

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



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 ρ.

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)

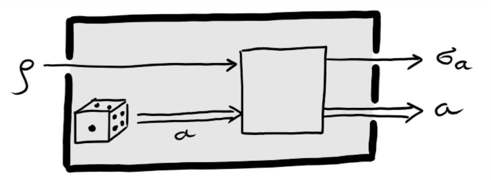
The information gained in a measurement resides in the way the outcomes are distributed.

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 ρ and all a
2. Suppose that there exist two states $\rho \neq \sigma$, 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 $\rho \neq \sigma$, 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 ρ. 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.

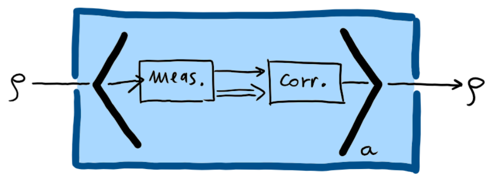
Definition (Stochastically Reversible Measurements)

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 ρ	there exist CPTP maps $\{\mathcal{R}_a\}_a$ such that $\sum_a (\mathcal{R}_a \circ \mathcal{E}_a)(\rho) = \rho$ for all ρ



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

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

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

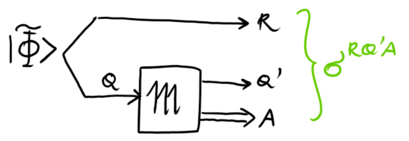
Two problems with the present formulation:

- 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

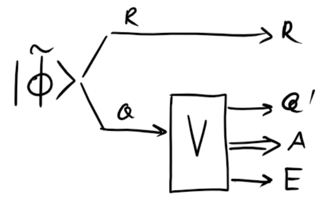
$$\delta(\mathcal{M}) \geq \iota(\mathcal{M})$$

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, $\delta(\mathcal{M}) = S(R) - S(Q'\mathbb{A}) + S(RQ'\mathbb{A}) = S(R) - S(RE_1E_2) + S(E_1E_2) = I(R;E_1E_2) = I(R;E_1\mathbb{A}) \geq I(R;\mathbb{A}) = \iota(\mathcal{M})$

□



The General Balance of Information

The global tradeoff is not able to account for the fact that it is possible to perfectly discriminate orthogonal quantum states without causing any disturbance.

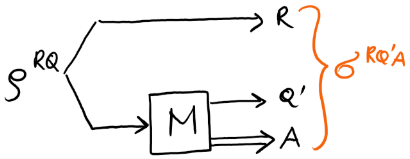
Some Clarifications

In the previous statement (global tradeoff), quantum disturbance and quantum information gain have been introduced as "global" parameters characterizing a given measurement apparatus.

For example, $\delta(\mathcal{M}) > 0$ simply means that *there are some states* that will be disturbed by the measurement. Analogously, $\iota(\mathcal{M}) > 0$ simply means that *there are some states* that will give rise to different outcome probability distributions.

Now we want to be more specific, and define both information and disturbance with respect to some *restricted* set of states, so to cover also the case of classical (commuting) states, which we know can be measured without disturbance.

Information Calculus



- R is the "reference," with respect to which "information" is about
- initially, the system Q carries $I(R;Q)$ bits of information
- after the measurement, the combined quantum+classical output $Q'\mathbb{A}$ carries $I(R;Q'\mathbb{A})$ bits
- by the chain rule of mutual information, this amount is divided as $I(R;Q'\mathbb{A}) = I(R;\mathbb{A}) + I(R;Q'|\mathbb{A})$
- here, $I(R;\mathbb{A})$ is the information carried by the outcome, while $I(R;Q'|\mathbb{A})$ is the information left in the system

Global Balance of Information

$$\underbrace{I(R;Q) - I(R;Q'\mathbb{A})}_{\text{net info-loss}} = \underbrace{I(R;Q) - I(R;Q'|\mathbb{A})}_{\text{system's info-loss}} - \underbrace{I(R;\mathbb{A})}_{\text{info-gain}}$$

Physical Interpretation

Global Balance of Information

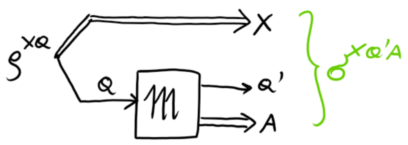
$$\underbrace{I(R;Q) - I(R;Q'\mathbb{A})}_{\text{net info-loss}} = \underbrace{I(R;Q) - I(R;Q'|\mathbb{A})}_{\text{system's info-loss}} - \underbrace{I(R;\mathbb{A})}_{\text{info-gain}}$$

- information never increases: by the data-processing theorem, the net loss is always non-negative, i.e., $I(R;Q) - I(R;Q'\mathbb{A}) \geq 0$
- the net informaton loss is an irreversibility parameter: as shown in [Junge et al, 2015],

$$\begin{aligned} & I(R;Q) - I(R;Q'\mathbb{A}) \\ & \geq -\log_2 \sup_{\{\mathcal{R}_a\}_a} \mathsf{F}\left\{\rho^{RQ}, \left[\mathsf{id} \otimes \sum_a (\mathcal{R}_a^{Q'} \circ \mathcal{E}_a^{Q})\right](\rho^{RQ})\right\} \end{aligned}$$

- in general, it is the net information loss, and not the information gain, to determine how much irreversible a measurement is

Example: The Case of a Classical Reference



- before the measurement: $\rho^{\mathbb{X}Q} = \sum_x p(x)|x\rangle\langle x|^{\mathbb{X}} \otimes \rho_x^Q$
- after the measurement: $\sigma^{\mathbb{X}Q'\mathbb{A}} = \sum_x p(x)|x\rangle\langle x|^{\mathbb{X}} \otimes \mathcal{E}_a^Q(\rho_x^Q) \otimes |a\rangle\langle a|^{\mathbb{A}}$
- the general balance of information in this case becomes

$$\underbrace{I(\mathbb{X};Q) - I(\mathbb{X};Q'\mathbb{A})}_{\text{net info-loss (disturbance)}} = \underbrace{I(\mathbb{X};Q) - I(\mathbb{X};Q'|\mathbb{A})}_{\text{system's info-loss}} - \underbrace{I(\mathbb{X};\mathbb{A})}_{\text{info-gain}} \geq 0$$

Example

Consider the fully classical situation where $\rho^{\mathbb{X}Q} = \sum_x p(x)|x\rangle\langle x|^{\mathbb{X}} \otimes |x\rangle\langle x|^Q$. Imagine a perfect measurement of x, i.e., $\mathcal{E}_{\bar{x}}(\bullet) = |\bar{x}\rangle\langle\bar{x}| \bullet |\bar{x}\rangle\langle\bar{x}|$. For such a measurement, the information gain is maximal, i.e., $I(\mathbb{X};\mathbb{A}) = H(\mathbb{X})$, even though the disturbance is zero!

When Does Information Gain Imply Disturbance?

- When the initial state ρ_{RQ} is pure, the following inequality holds:

$$\underbrace{I(R;Q) - I(R;Q'|\mathbb{A})}_{\text{system's info-loss}} \geq 2\,\underbrace{I(R;\mathbb{A})}_{\text{info-gain}}$$

- When the above equation holds, then we have the information-disturbance tradeoff

$$\underbrace{I(R;Q) - I(R;Q'\mathbb{A})}_{\text{net info-loss (disturbance)}} \geq \underbrace{I(R;\mathbb{A})}_{\text{info-gain}}$$

- if ρ^{RQ} is pure, hence, information gain necessarily requires irreversibility; otherwise, this need not be true

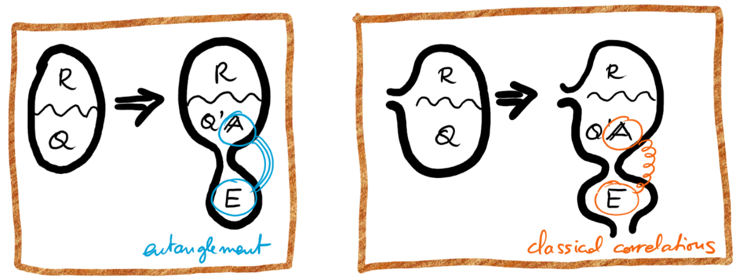
Remark. When ρ^{RQ} is the maximally entangled state, the information gain becomes the (global) quantum information gain $\iota(\mathcal{M})$ and the net information loss becomes that (global) quantum disturbance $\delta(\mathcal{M})$.

The Case of Pure States

As we saw, the information-disturbance tradeoff holds whenever

$$\underbrace{I(R;Q) - I(R;Q'|\mathbb{A})}_{\text{system's info-loss}} \geq 2\,\underbrace{I(R;\mathbb{A})}_{\text{info-gain}},$$

which is true if the initial system+reference state is pure. How to "visualize" this?



In a closed system (pure state) correlations between $\mathbb{A}$ and E must be quantum correlations (entanglement). Not so if the system is open from the start (mixed state).

An Interesting Byproduct

In proving the general tradeoff, we obtained a stronger data-processing inequality for quantum measurements, valid when the entire measurement $\{\mathcal{E}_a\}_a$, and not only the corresponding POVM, is known.

- simple data-processing inequality: $I(R;Q) \geq I(R;\mathbb{A})$
- stronger form: $I(R;Q) - I(R;Q'|\mathbb{A}) \geq I(R;\mathbb{A})$
- strongest form: if RQ is in a pure state, $\frac{1}{2}[I(R;Q) - I(R;Q'|\mathbb{A})] \geq I(R;\mathbb{A})$

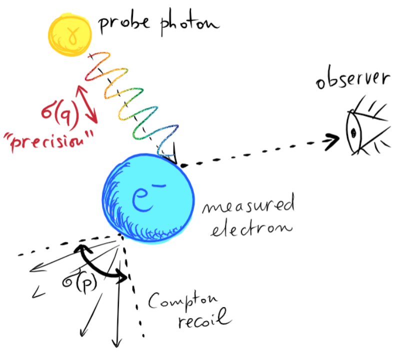
In the case of a classical reference, we obtain a refined Holevo bound.

- simple Holevo bound: $I(\mathbb{X};Q) \geq I(\mathbb{X};\mathbb{A})$
- stronger form: $I(\mathbb{X};Q) - I(\mathbb{X};Q'|\mathbb{A}) \geq I(\mathbb{X};\mathbb{A})$
- (no strongest form, because the initial state cannot be pure)

See also [Schumacher, Westmoreland, Wootters, PRL, 1996].

Heisenberg's Two-Observable Formulation

Heisenberg's γ-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. Notice that here the electron's state is not explicitly mentioned. Heisenberg's formulation of disturbance is with respect two "properties" (i.e., dynamical variables, observables, etc) of the electron.

Modern Approaches to Heisenberg's Problem

- It is well known now that Heisenberg's relation $\sigma(p)\sigma(q) \geq \frac{\hbar}{2}$, and its generalization due to Robertson $\sigma(A)\sigma(B) \geq \frac{1}{2}\langle[A,B]\rangle$, should be interpreted as posing a constraint on the preparation of quantum states, rather than a constraint on quantum measurements*
- To save Heisenberg's original intuition, there are at present two main approaches:
  1. state-dependent approach (Ozawa)
  2. state-independent approach (Busch, Lahti, and Werner)
- In what follows, we will see an information-theoretic formulation of the state-independent approach

*The quantity $\sigma(A)$ is defined as $\sigma(A) = \sqrt{\mathrm{Var}_\psi(A)} = \sqrt{\langle\psi|(A^2 - \langle A\rangle^2)|\psi\rangle}$. Hence, the inequality $\sigma(A)\sigma(B) \geq \frac{1}{2}\langle[A,B]\rangle$ is usually interpreted as saying that it is impossible to prepare a state $|\psi\rangle$ that is simultaneously sharp in both observables A and B, if $\langle\psi|[A,B]|\psi\rangle$ is not zero.

Heisenberg's Two-Observable Problem

On a d-dimensional Hilbert space, consider two non-degenerate observables $X = \sum_{x=1}^{d} \xi_x |\psi_x\rangle\langle\psi_x|$ and $Z = \sum_{z=1}^{d} \zeta_z |\varphi_z\rangle\langle\varphi_z|$.

Questions: given a measurement $\mathcal{M} = \{\mathcal{E}_a\}_a$, how much information about X is $\mathcal{M}$ able to extract? How much does $\mathcal{M}$ disturb observable Y? How are such quantities related?

X-Information (informal definition)

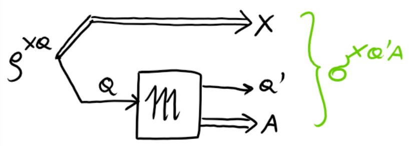
Imagine to input each eigenstate $|\psi_x\rangle$ of X in $\mathcal{M}$, with uniform a priori probability $p(x) = \frac{1}{d}$. The information provided by $\mathcal{M}$ is measured by how much the outcome index a is correlated with the input label x.

Y-Disturbance (informal definition)

Imagine to input each eigenstate $|\varphi_z\rangle$ of Z in $\mathcal{M}$, with uniform a priori probability $p(z) = \frac{1}{d}$. The disturbance caused by $\mathcal{M}$ is measured by how much irreversible the action of $\mathcal{M}$ is on the eigenstates of Z.

The above two definitions can be formalized using the concepts of information gain and net information loss introduced before.

Information about X: The X-Error



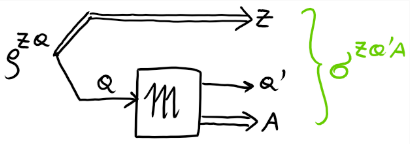
- having in mind the above setting, let us assume:
  $\rho^{\mathbb{X}Q} = \frac{1}{d}\sum_{x=1}^{d} |x\rangle\langle x|^{\mathbb{X}} \otimes |\psi_x\rangle\langle\psi_x|^{Q}$
- after the measurement we have:
  $\sigma^{\mathbb{X}Q'\mathbb{A}} = \frac{1}{d}\sum_{x=1}^{d} |x\rangle\langle x|^{\mathbb{X}} \otimes \mathcal{E}_a^{Q}(|\psi_x\rangle\langle\psi_x|^{Q}) \otimes |a\rangle\langle a|^{\mathbb{A}}$

Definition (X-Error)

The error that measurement $\mathcal{M} = \{\mathcal{E}_a\}_a$ does when used to measure the observable X is defined as the difference between perfect information, i.e., $\log d$, and the information gain $I(\mathbb{X};\mathbb{A})$, computed with respect to $\sigma^{\mathbb{X}Q'\mathbb{A}}$ above. In formula,

$$\epsilon_X(\mathcal{M}) \triangleq \log d - \iota_X(\mathcal{M}) = H(\mathbb{X}|\mathbb{A}) \ .$$

Disturbance on Z: The Z-Disturbance



- let us now assume: $\rho^{\mathbb{Z}Q} = \frac{1}{d}\sum_{z=1}^{d} |z\rangle\langle z|^{\mathbb{Z}} \otimes |\varphi_z\rangle\langle\varphi_z|^{Q}$
- after the measurement we have: $\sigma^{\mathbb{Z}Q'\mathbb{A}} = \frac{1}{d}\sum_{z=1}^{d} |z\rangle\langle z|^{\mathbb{Z}} \otimes \mathcal{E}_a^{Q}(|\varphi_z\rangle\langle\varphi_z|^{Q}) \otimes |a\rangle\langle a|^{\mathbb{A}}$

Definition (Z-Disturbance)

The disturbance that measurement $\mathcal{M} = \{\mathcal{E}_a\}_a$ causes on the observable Z is defined as the net information loss

$$\delta_Z(\mathcal{M}) \triangleq I(\mathbb{Z};Q) - I(\mathbb{Z};Q'\mathbb{A}) \ ,$$

computed with respect to $\sigma^{\mathbb{Z}Q'\mathbb{A}}$ above.

The Error-Disturbance Tradeoff Relation

Theorem

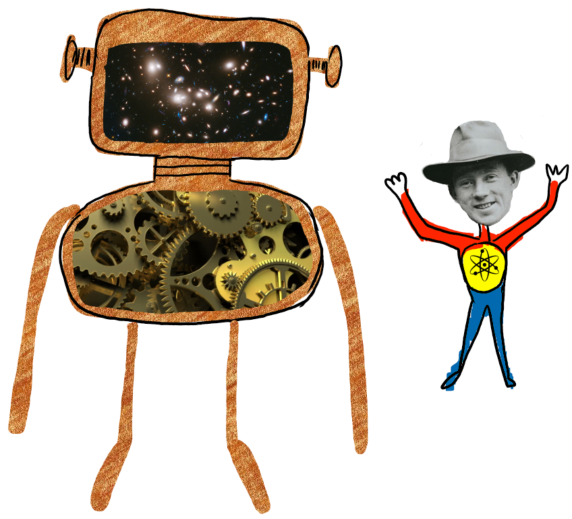
For any pair of non-degenerate observables $X = \sum_x \xi_x |\psi_x\rangle\langle\psi_x|$ *and* $Z = \sum_z \zeta_z |\varphi_z\rangle\langle\varphi_z|$*, and for any measurement* $\mathcal{M} = \{\mathcal{E}_a\}_a$*, we have*

$$\epsilon_X(\mathcal{M}) + \delta_Z(\mathcal{M}) \geq -\log c \, ,$$

where $c \triangleq \max_{x,z} |\langle\psi_x|\varphi_z\rangle|^2$*.*

- any measurement can be evaluated against two observables X and Z, even measurements that have a number of outcomes different from d
- the bound becomes trivial, i.e., $c = 1$, if and only if X and Z have one common eigenstate
- the proof of the relation above does not follow directly from any of the information-disturbance tradeoff relations we showed before; formally, it is a consequence of Maassen's and Uffink's entropic uncertainty relations, however, the interpretation is very different
- the error $\epsilon_X(\mathcal{M})$ measures how correlated the outcome of the measurement is with the eigenvalues of X; however, the actual numerical values of the eigenvalues do not play any role here (contrarily to what happens with the definitions involving variance-like measures)
- the disturbance $\delta_Z(\mathcal{M})$, as we saw before, is directly related with the possibility of reversing (i.e., correcting) the action of $\mathcal{M}$ on the eigenstates of Z

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



Some References

- H.J. Groenewold: *Information Gain in Quantal Measurements*. Int. J. Theor. Phys. **4**, 327 (1971).
- A. Barchielli, G. Lupieri: *Quantum Measurements and Entropic Bounds*. Quantum Inf. Comput. **6**, 16 (2006)
- L. Maccone: *Entropic Information-Disturbance Tradeoff*. Europhys. Lett. **77**, 40002 (2007).
- F. Buscemi, M. Hayashi, M. Horodecki: *Global Information Balance in Quantum Measurements*. Phys. Rev. Lett. **100**, 210504 (2009)
- F. Buscemi, M.J.W. Hall, M. Ozawa, M.M. Wilde: *Noise and Disturbance in Quantum Measurements: An Information-Theoretic Approach*. Phys. Rev. Lett. **112**, 050401 (2014)