Entropy and information gain in quantum continual measurements

Alberto Barchielli
Politecnico di Milano, Dipartimento di Matematica,
Piazza Leonardo da Vinci 32, I-20133 Milano, Italy;
also Istituto Nazionale di Fisica Nucleare, Sezione di Milano.
E-mail: barchielli@mate.polimi.it

1 Introduction

The theory of measurements continuous in time in quantum mechanics (quantum continual measurements) has been formulated by using the notions of instrument, positive operator valued (POV) measure, etc. [1,2], by using quantum stochastic differential equations [3,4] and by using classical stochastic differential equations (SDE’s) for vectors in Hilbert spaces or for trace-class operators [5,6,7,8]. In the same times Ozawa made developments in the theory of instruments [9,10] and introduced the related notions of a posteriori states [11] and of information gain [12].

In Section 2 we introduce a simple class of SDE’s relevant to the theory of continual measurements and we recall how they are related to instruments and a posteriori states and, so, to the general formulation of quantum mechanics [13]. In Section 3 we shall introduce and use the notion of information gain and the other results of paper [12] inside the theory of continual measurements.

2 Stochastic differential equations and instruments

Let $H$ be a separable complex Hilbert space, associated to the quantum system of interest. Let us denote by $B(H)$ the space of bounded linear operators on $H$ and by $T(H)$ the trace-class on $H$, i.e. $T(H) = \{ \rho \in B(H) : \| \rho \| = \text{Tr} \{ \sqrt{\rho^* \rho} \} < \infty \}$. Let $S(H) \subset T(H)$ be the set of all statistical operators (states) on $H$. Commutators and anticommutators are denoted by $[,]$ and $\{,\}$, respectively.

Let $H, L_j, S_h, j, h = 1, 2, \ldots$, be bounded operators on $H$ such that $H = H^\dagger, \sum_{j=1}^\infty L_j^\dagger L_j$ and $\sum_{h=1}^\infty S_h^\dagger S_h$ are strongly convergent in $B(H)$. Let $J_k$ be a bounded linear map on $T(H)$ such that its adjoint $J_k^\dagger$ is a normal, completely positive map on $B(H)$ and $\sum_{k=1}^\infty J_k^\dagger[I]$ is strongly convergent to a bounded
operator. Then, we introduce the following operators on $\mathcal{T}(\mathcal{H})$:

$$
L_0[\rho] = -i[H, \rho] + \sum_{j=1}^{\infty} \left( L_j \rho L_j^\dagger - \frac{1}{2} \left\{ L_j^\dagger L_j , \rho \right\} \right) + \sum_{k=1}^{\infty} \left( J_k[\rho] - \frac{1}{2} \left\{ J_k[^\dagger], \rho \right\} \right),
$$

(1)

$$
L_1[\rho] = \sum_{h=1}^{\infty} \left( S_h \rho S_h^\dagger - \frac{1}{2} \left\{ S_h^\dagger S_h , \rho \right\} \right),
$$

(2)

$$
L = L_0 + L_1.
$$

(3)

The adjoint operators of $L, L_0, L_1$ are generators of norm-continuous quantum dynamical semigroups [14, 15].

Let us now consider the following linear SDE (in the sense of Itô) for trace-class operators:

$$
d\sigma_t = L[\sigma_t] \, dt + \sum_{j=1}^{\infty} \left( \tilde{L}_j(t) \sigma_t + \sigma_t \tilde{L}_j(t)^\dagger \right) d\tilde{W}_j(t) + \sum_{k=1}^{\infty} \left( \frac{1}{\lambda_k} J_k[\sigma_t] - \sigma_t \right) (dN_k(t) - \lambda_k \, dt);
$$

(4)

the initial condition is $\sigma_0 = \rho \in \mathcal{S}(\mathcal{H})$ (a non-random state) and we have set

$$
\tilde{L}_j(t) = e^{i\omega_j t} L_j , \quad \omega_j \in \mathbb{R}.
$$

(5)

The processes $\tilde{W}_j(t)$ are independent standard Wiener processes, the $N_k(t)$ are independent Poisson processes of intensity $\lambda_k > 0$, which are also independent of the Wiener processes; we assume $\sum_k \lambda_k < +\infty$.

These processes are realized in a probability space $(\Omega, \mathcal{F}, Q)$; the sample space $\Omega$ is, roughly speaking, the set of possible trajectories for the processes $\tilde{W}_j$, $N_k$, the event space $\mathcal{F}$ is the $\sigma$-algebra of sets of trajectories to which a probability can be given and $Q$ is the probability law under which $\tilde{W}_j, N_k$ are independent Wiener and Poisson processes. Moreover, let $\mathcal{F}_t$ be the collection of events which are specified by giving conditions involving times only in the interval $[0, t]$. We also ask $\mathcal{F} = \mathcal{F}_\infty$. In mathematical terms the $\mathcal{W}_j$, $N_k$ are canonical Wiener and Poisson processes, $\{\mathcal{F}_t, t \geq 0\}$ is their natural filtration and $\mathcal{F} = \bigcup_{t \geq 0} \mathcal{F}_t$. Finally, let us denote by $E_Q$ the expectation with respect to the probability $Q$, i.e. $E_Q[A] = \int_\Omega A(\omega) Q(\, d\omega)$.

For every $F \in \mathcal{F}_t$ and every initial condition $\rho \in \mathcal{S}(\mathcal{H})$, let us set

$$
\mathcal{I}_t(F)[\rho] = E_Q[1_F \sigma_t] = \int_F \sigma_t(\omega) Q(\, d\omega);
$$

(6)

$1_F$ is the indicator function of the set $F$, i.e. $1_F(\omega) = 1$ if $\omega \in F$ and $1_F(\omega) = 0$ if $\omega \notin F$. The map $\mathcal{I}_t$ turns out to be a (completely positive) instrument [10].

2
with value space \((\Omega, \mathcal{F}_t)\) and \(\mathcal{I}_t(\cdot)^*[\mathbb{I}]\) is the associated POVM measure. Then we set, \(\forall F \in \mathcal{F}_t\),
\[
P_\rho(F) = \text{Tr}\{\mathcal{I}_t(F)^*[\mathbb{I}]\rho\} = E_Q[\|\sigma_t\| 1_F].
\] (7)

The important point in this formula is that \(\|\sigma_t\|\) is a Q-martingale and this implies that the time dependent probability measures on the r.h.s. are consistent and define a unique probability \(P_\rho\) on \((\Omega, \mathcal{F})\).

The interpretation of eqs. (6) and (7) is that \(\{\mathcal{I}_t, t \geq 0\}\) is the family of instruments describing the continual measurement, the processes \(\tilde{W}_j, N_k\) represent the output of this measurement and \(P_\rho\) is the physical probability law of the output.

From eq. (6) it follows that
\[
\eta_t = \mathcal{I}_t(\Omega)[\rho] = E_Q[\sigma_t]
\] (8)
is the state to be attributed to the system at time \(t\) if the output of the measurement is not taken into account or not known; it can be called the a priori state at time \(t\). It turns out that the a priori states satisfy the master equation
\[
\frac{d}{dt} \eta_t = \mathcal{L}[\eta_t], \quad \eta_0 = \rho.
\] (9)

If we introduce the random states
\[
\rho_t = \frac{\sigma_t}{\|\sigma_t\|},
\] (10)
then we have, \(\forall F \in \mathcal{F}_t\),
\[
\mathcal{I}_t(F)[\rho] = E_Q[1_F \sigma_t] = E_{P_\rho}[1_F \frac{\sigma_t}{\|\sigma_t\|}] = \int_F \rho_t(\omega) P_\rho(d\omega).\] (11)

According to [11], \(\rho_t(\omega)\) is a family of a posteriori states for the instrument \(\mathcal{I}_t\) and the initial state \(\rho\), i.e. \(\rho_t(\omega)\) is the state to be attributed to the system at time \(t\) when the trajectory \(\omega\) of the output is known, up to time \(t\). Note that \(\eta_t = E_Q[\sigma_t] = E_{P_\rho}[\rho_t]\).

By using Itô’s calculus, we find that the a posteriori states satisfy the non-linear SDE
\[
d\rho_t = \mathcal{L}[\rho_t] dt + \sum_{j=1}^{\infty} \left[ \frac{1}{\nu_k(t)} J_k[\rho_t] - \rho_t \right] \frac{dN_k(t) - \nu_k(t) dt}{\nu_k(t)},
\] (12)
where
\[
W_j(t) = \tilde{W}_j(t) - \int_0^t m_j(s) ds,
\] (13)
\[
m_j(t) = \text{Tr}\{\rho_t - \left( L_j(t) + \tilde{L}_j(t) \right)^\dagger\}, \quad \nu_k(t) = \text{Tr}\{\rho_t - J_k[\mathbb{I}]\}.\] (14)
Under the physical probability law \( P_\rho \), the processes \( W_j(t) \) are independent standard Wiener processes and the \( N_k(t) \) are counting processes with stochastic intensity \( \nu_k(t) \). In eq. [12] the sum in the jump term is only on the set where the stochastic intensity \( \nu_k(t) \) is different from zero.

Formulae for the moments of the output can be obtained by the technique of the characteristic operator [2, 3, 4]. Let \( h_\alpha \) be real test functions in a suitable space; we define the characteristic operator \( \mathcal{G} \) by

\[
\mathcal{G}_t(h)[\rho] = \mathbb{E}_{P_\rho} \left[ \exp \left\{ i \sum_j \int_0^t h_{j1}(s) d\tilde{W}_j(s) \right\} \rho_t \right],
\]

then, \( \text{Tr} \{ \mathcal{G}_t(h)[\rho] \} \) is the characteristic functional of the output up to time \( t \) (the Fourier transform of \( P_\rho \) restricted to \( \mathcal{F}_t \)). By Itô’s calculus we obtain

\[
\frac{d}{dt} \mathcal{G}_t(h)[\rho] = \mathcal{K}_t(h) \circ \mathcal{G}_t(h)[\rho],
\]

\[
\mathcal{K}_t(h)[\rho] = \mathcal{L}[\rho] + i \sum_j h_{j1}(t) \left[ \tilde{L}_j(t) \rho + \rho \tilde{L}_j(t)^\dagger \right] - \frac{1}{2} \sum_j h_{j1}(t)^2 \rho + \sum_k \{ \exp[ih_{k2}(t)] - 1 \} J_k[\rho].
\]

All the moments can be obtained by functional differentiation of the characteristic functional. In particular, the mean values are expressed in terms of the a priori states as

\[
\mathbb{E}_{P_\rho} [\tilde{W}_j(t)] = \int_0^t \mathbb{E}_{P_\rho} [m_j(s)] ds, \quad \mathbb{E}_{P_\rho} [N_k(t)] = \int_0^t \mathbb{E}_{P_\rho} [\nu_k(s)] ds,
\]

\[
\mathbb{E}_{P_\rho} [m_j(s)] = \text{Tr} \left\{ \eta_s \left( \tilde{L}_j(s) + \tilde{L}_j(s)^\dagger \right) \right\}, \quad \mathbb{E}_{P_\rho} [\nu_k(s)] = \text{Tr} \{ J_k[\eta_s] \},
\]

and the second moments are given by

\[
\mathbb{E}_{P_\rho} [X_{j\alpha}(t) X_{j\beta}(s)] = \delta_{ij} \delta_{\alpha\beta} \int_0^{\min\{t,s\}} d\tau \left( \delta_{\alpha 1} + \delta_{\alpha 2} \text{Tr} \{ J_i[\eta_s] \} \right)
\]

\[
+ \int_0^t d\tau_1 \int_0^{\min\{s,\tau_1\}} d\tau_2 \text{Tr} \left\{ A_{j\alpha}(\tau_1) \circ e^{L(\tau_1-\tau_2)} \circ A_{j\beta}(\tau_2)[\eta_{\tau_2}] \right\}
\]

\[
+ \int_0^s d\tau_2 \int_0^{\min\{t,\tau_2\}} d\tau_1 \text{Tr} \left\{ A_{j\beta}(\tau_2) \circ e^{L(\tau_2-\tau_1)} \circ A_{j\alpha}(\tau_1)[\eta_{\tau_1}] \right\},
\]

where \( X_{j1}(t) = \tilde{W}_j(t) \), \( X_{j2}(t) = N_j(t) \), \( A_{j1}(t)[\rho] = \tilde{L}_j(t) \rho + \rho \tilde{L}_j(t)^\dagger \), \( A_{j2}(t) = J_j \).
The class of SDE's presented here is a particular case of the one studied in [13], and, while not so general, it contains the main detection schemes found in quantum optics [3]; also the chosen time-dependence is natural for some systems typical of quantum optics under the so called heterodyne/homodyne detection scheme.

3 Entropy and information gain

In [12] a measurement is called quasi-complete if the \textit{a posteriori} states are pure for every pure initial state and it is called complete if the \textit{a posteriori} states are pure for every (pure or mixed) initial state. So, we call \textit{quasi-complete} the continual measurement of Section 2 if the \textit{a posteriori} states \( \rho_t \) are pure (\( P_\rho \)-almost surely) for all \( t \) and for all pure initial conditions \( \rho \). In [13], we proved that

\textbf{Theorem 1} The continual measurement of Section 2 is quasi-complete if and only if \( L_1 = 0 \) and \( \frac{J_k[\rho]}{\text{Tr}\{J_k[\rho]\}} \) is a pure state for every \( k \) and for every pure state \( \rho \). In this case there exists a partition \( A_1, A_2 \) of the integer numbers such that for some \( R_k \in \mathcal{B}(\mathcal{H}) \) and for some monodimensional projection \( P_k \) we can write \( J_k[\rho] = R_k \rho R_k^\dagger \), for \( k \in A_1 \), \( J_k[\rho] = \text{Tr}\{\rho J_k^*[\mathbb{I}]\} P_k \), for \( k \in A_2 \).

Our continual measurement can not be complete in the sense of [12] for a fixed time; however, it can be “asymptotically complete”. Examples of this behaviour in the case of linear systems are given in [19]. In [18], we proved that

\textbf{Theorem 2} Let the continual measurement of Section 3 be quasi-complete and let \( \mathcal{H} \) be finite-dimensional. If for every time \( t \) it does not exist a bidimensional projection \( P_t \) such that, \( \forall j, k, P_t \left( \bar{L}_j(t) + \bar{L}_j(t)^\dagger \right) P_t = z_j(t) P_t, P_t J_k^*[\mathbb{I}] P_t = q_k(t) P_t \) for some complex numbers \( z_j(t) \) and \( q_k(t) \), then eq. (12) maps asymptotically, for \( t \to \infty \), mixed states into pure ones, in the sense that for every initial condition \( \rho \) we have \( P_\rho \)-almost surely \( \lim_{t \to \infty} \text{Tr}\{\rho_t (\mathbb{I} - \rho_t)\} = 0 \).

The proof of the theorems above is based on the study of the \textit{a posteriori} linear entropy (or purity) \( \text{Tr}\{\rho_t (\mathbb{I} - \rho_t)\} \) and of its mean value. However, physically more interesting quantities are the von Neumann entropy and the relative entropy: for \( x, y \in \mathcal{S}(\mathcal{H}) \), \( S[x] = -\text{Tr}\{x \ln x\} \geq 0 \), \( S[x|y] = \text{Tr}\{x \ln x - x \ln y\} \geq 0 \) (they can also diverge) [16]. In our case we have the initial state \( \rho = \rho_0 = \sigma_0 = \eta_0 \) and the \textit{initial entropy} \( S[\rho] \), the \textit{a priori} state \( \eta_t \) and the \textit{a priori entropy} \( S[\eta] \), the \textit{a posteriori} states \( \rho_t \) and the mean \textit{a posteriori entropy}

\[ E_{P_\rho} \left[ S[\rho_t] \right] = E_Q \left[ \|\sigma_t\| \ln \|\sigma_t\| - \text{Tr}\{\sigma_t \ln \sigma_t\} \right]. \tag{18} \]

By some direct computations, we obtain a first relation among these quantities:

\[ S[\eta] - E_{P_\rho} \left[ S[\rho_t] \right] = E_{P_\rho} \left[ S[\rho_t|\eta_t] \right] \geq 0. \tag{19} \]
Following [12], we can also introduce the amount of information of the continual measurement
\[ I[\rho; t] = S[\rho] - \mathbb{E}_{P_\rho}[S[\rho_t]] \]  
and the classical amount of information. To introduce this last quantity we need some notations. Let us set \( P_\rho(d\omega; t) = \|\sigma_t(\omega)\|Q(d\omega) \), let \( \rho = \sum_\alpha w_\alpha \rho_\alpha \) be the orthogonal decomposition of \( \rho \) into pure states and \( P_\rho, \sigma_t^\alpha, \rho_t^\alpha, \eta_t^\alpha, m_j^\alpha(t), \nu_k^\alpha(t) \) be defined starting from \( \rho_\alpha \) as \( P_\rho, \sigma_t, \rho_t, \eta_t, m_j(t), \nu_k(t) \) are defined starting from \( \rho \). Then, the classical amount of information of the continual measurement is defined by
\[
\text{c-}I[\rho; t] = \sum_\alpha w_\alpha \int_{\Omega} \ln \left( \frac{P_{\rho_\alpha}(d\omega; t)}{P_\rho(d\omega; t)} \right) P_{\rho_\alpha}(d\omega; t)
\]
\[
= \sum_\alpha w_\alpha \mathbb{E}_{P_{\rho_\alpha}} \left[ \ln \| \sigma_t^{\alpha} \| \right]
\]
\[
= \mathbb{E}_Q \left[ \sum_\alpha w_\alpha \| \sigma_t^{\alpha} \| \ln \| \sigma_t^{\alpha} \| - \| \sigma_t \| \ln \| \sigma_t \| \right]. \tag{21}
\]
By classical arguments, c-\( I[\rho; t] \) is always positive [12]: c-\( I[\rho; t] \) \( \geq 0 \), \( \forall t \geq 0 \), \( \forall \rho \in \mathcal{S}(\mathcal{H}) \). Obviously, we have \( I[\rho; t] \leq S[\rho], I[\rho; 0] = 0 \), c-\( I[\rho; 0] = 0 \). If it exists an equilibrium state \( \eta_{\text{eq}} \) \( (\mathcal{L}[\eta_{\text{eq}}] = 0) \), by [13] we have also \( I[\eta_{\text{eq}}; t] \) \( \geq 0 \).

**Theorem 3**

The classical amount of information of the continual measurement of Section 2 is non-decreasing in time and
\[
\frac{d}{dt} \text{c-}I[\rho; t] = \sum_\alpha w_\alpha \mathbb{E}_{P_{\rho_\alpha}} \left[ \frac{1}{2} \sum_j m_j^\alpha(t)^2 + \sum_k \nu_k^\alpha(t) \ln \nu_k^\alpha(t) \right]
\]
\[
- \mathbb{E}_{P_\rho} \left[ \frac{1}{2} \sum_j m_j(t)^2 + \sum_k \nu_k(t) \ln \nu_k(t) \right]
\]
\[
= \sum_\alpha w_\alpha \mathbb{E}_{P_{\rho_\alpha}} \left[ \frac{1}{2} \sum_j (m_j^\alpha(t) - m_j(t))^2 \right]
\]
\[
+ \sum_k \nu_k(t) \left( 1 - \frac{\nu_k^\alpha(t)}{\nu_k(t)} + \frac{\nu_k^\alpha(t)}{\nu_k(t)} \ln \frac{\nu_k^\alpha(t)}{\nu_k(t)} \right) \geq 0. \tag{22}
\]
To prove this theorem one has to differentiate the last expression in (21) and to use the relationships among \( Q, P_\rho, P_{\rho_\alpha} \).

For quasi-complete measurements the information gain \( I[\rho; t] \) has a nice behaviour.

**Theorem 4**
The continual measurement of Section 2 is quasi-complete if and only if the amount of information \( I[\rho; t] \) is non-negative for any \( \rho \in \mathcal{S}(\mathcal{H}) \) with \( S[\rho] < +\infty \) and any \( t \geq 0 \). Moreover, if it is quasi-complete, we have
\[ I[\rho; t] \geq I[\rho; s] \geq I[\rho; t] \geq 0, \quad I[\rho; t] \geq I[\rho; s] \text{ for any } t, \text{ any } s < t \text{ and any state } \rho \text{ with } S[\rho] < +\infty. \]

**Proof.** All the statements but the last one are a particularization of Theorems 1 and 2 of [12] to our case. The last statement needs the use of conditional expectations. We have \( I[\rho; t] - I[\rho; s] = \mathbb{E}_{P_\rho}[S[\rho_s] - \mathbb{E}_{P_\rho}[S[\rho_t]|F_s]]; \) by (12) \( S[\rho_s] - \mathbb{E}_{P_\rho}[S[\rho_t]|F_s] \) is the amount of information at time \( t \) when the initial time is \( s \) and the initial state is \( \rho_s \) and, so, it is non-negative for a quasi-complete measurement. \( \square \)

Finally, if \( \mathcal{H} \) is finite-dimensional, the vanishing of the purity implies the vanishing of the entropy; therefore, we have the asymptotic completeness also in the sense of the vanishing of the entropy:

\[
\text{The hypotheses of Theorem 2 imply also that } \lim_{t \to +\infty} S[\rho_t] = 0, \text{ } P_\rho-\text{almost surely, and } \lim_{t \to +\infty} I[\rho; t] = S[\rho].
\]

**References**

[1] E. B. Davies, *Quantum Theory of Open Systems* (Academic Press, London, 1976).

[2] A. Barchielli, L. Lanz, G. M. Prosperi, Statistics of continuous trajectories in quantum mechanics: Operation valued stochastic processes, Found. Phys. 13 (1983) 779–812.

[3] A. Barchielli, G. Lupieri, Quantum stochastic calculus, operation valued stochastic processes and continual measurements in quantum mechanics, J. Math. Phys. 26 (1985) 2222–2230.

[4] A. Barchielli, Direct and heterodyne detection and other applications of quantum stochastic calculus to quantum optics, Quantum Opt. 2 (1990) 423–441.

[5] V. P. Belavkin, Nondemolition measurements, nonlinear filtering and dynamic programming of quantum stochastic processes. In A. Blaquière (Ed.), *Modelling and Control of Systems*, Lecture Notes in Control and Information Sciences 121 (Springer, Berlin, 1988) pp. 245–265.

[6] V. P. Belavkin, A new wave equation for a continuous nondemolition measurement, Phys. Lett. A 140 (1989) 355–358.

[7] V. P. Belavkin, A continuous counting observation and posterior quantum dynamics, J. Phys. A: Math. Gen. 22 (1989) L1109–L1114.

[8] A. Barchielli, V. P. Belavkin, Measurements continuous in time and a posteriori states in quantum mechanics, J. Phys. A: Math. Gen. 24 (1991) 1495–1514.
[9] M. Ozawa, *Quantum measuring processes of continuous observables*, J. Math. Phys. **25** (1984) 79–87.

[10] M. Ozawa, *Mathematical characterizations of measurement statistics*. In V. P. Belavkin, O. Hirota, R. L. Hudson (Eds.), *Quantum Communications and Measurement (Nottingham, 1994)* (Plenum, New York, 1995) pp. 109–117.

[11] M. Ozawa, *Conditional probability and a posteriori states in quantum mechanics*, Publ. Res. Inst. Math. Sc. Kyoto Univ. **21** (1985) 279–295.

[12] M. Ozawa, *On information gain by quantum measurements of continuous observables*, J. Math. Phys. **27** (1986) 759–763.

[13] K. Kraus, *States, Effects and Operations*, Lecture Notes in Physics **190** (Springer, Berlin, 1980).

[14] G. Lindblad, *On the generators of quantum dynamical semigroups*, Commun. Math. Phys. **48** (1976) 119–130.

[15] R. Alicki, *General theory and applications to unstable particles*. In: R. Alicki, K. Lendi, *Quantum Dynamical Semigroups and Applications*, Lecture Notes in Physics **286** (Springer, Berlin, 1987) pp. 1–94.

[16] A. Barchielli, A. S. Holevo, *Constructing quantum processes via classical stochastic calculus*, Stochastic Process. Appl. **58** (1995) 293–317.

[17] A. Barchielli, A. M. Paganoni, *Detection theory in quantum optics: stochastic representation*, Quantum Semiclass. Opt. **8** (1996) 133–156.

[18] A. Barchielli, A. M. Paganoni, *On the asymptotic behaviour of some stochastic differential equations for quantum states*, preprint of the Mathematical Department n. 420/P, July 2000.

[19] A. C. Doherty, S. M. Tan, A. S. Parkins, D. F. Walls, *State determination in continuous measurements*, Phys. Rev. A **60** (1999) 2380–2392.