

The Observer Effect in Estimation with Physical Communication Constraints [★]

Henrik Sandberg* Jean-Charles Delvenne**

* *Automatic Control Lab, School of Electrical Engineering, KTH Royal
Institute of Technology, 100 44 Stockholm, Sweden. (e-mail:
hsan@ee.kth.se).*

** *Department of Applied Mathematics, Université catholique de
Lowain, 1348 Lowain-la-Neuve, Belgium. (e-mail:
jean-charles.delvenne@uclouvain.be)*

Abstract: In this paper, we derive physical limitations on measurements of linear systems. We show that there is a trade-off between how precise a measurement is and how much we must perturb the measured system. This effect, very different from Heisenberg's inequality, originates in the thermal noise of the communication line between the measured system and the measuring device. This observer effect is shown to be relevant for small systems as found in nanoscience and microbiology.

Keywords: Estimation theory; Limits; Physical models; Communication channels; Optimal filtering

1. INTRODUCTION

With the tremendous development of communication technologies in the recent decades, there has been a great amount of research developing a control theory that is consistent with the communication constraints occurring when plant and controller interact through lines of communication. As a sample of the communication constraints that have received the most attention, we cite finite-capacity digital channels (e.g., Nair et al. (2004); Martins et al. (2007)), delays (e.g., Sahai and Mitter (2006)) or packet drops (e.g., Sinopoli et al. (2003)). Although this is certainly a very incomplete review of the field, it is fair to say that in our knowledge, all communication constraints considered so far arise from technological limitations, which would disappear in the limit of large bandwidth, small delays, high quality of service, etc. In this paper we take a different direction. We look at an estimation problem where the limitation comes from fundamental laws of statistical physics, and only vanish when the temperature of the transmission medium goes to zero. As a consequence, both the results and the proofs are incomparable to previous results of the field of communication constraints in control.

Specifically, we introduce a new kind of limitation due to communication. We show that analog transmission media, such as coaxial cables or even simple elastic mechanical

beams, also introduce limitations on the efficiency of control and estimation. Namely, we show that estimating the state of a system through a transmission media cannot be done without perturbing the measured system. We quantify the trade-off between the measurement accuracy and the corresponding perturbation on the measured plant. Since the limitations are derived using a basic physical assumption, conservation of energy, we call these limitations *physical limitations*. In quantum mechanics, it is well known that a measurement causes a disturbance to the measured system. There is an *observer effect*. It is also known that there are observer effects in classical physics: To measure for example the pressure or the temperature of a system, we need to connect the system to a meter which slightly changes the dynamics of the measured system. But to derive fundamental limitations on this (classical) effect does not seem to have been a much studied problem. In Barnes and Silverman (1934); McCombie (1953), steps in this direction are taken. But at that time, optimal filtering theory was not as well developed as it is now. In this paper, we combine modeling that respects basic physical laws (conservation of energy) with filtering theory to bound a classical observer effect.

Let us detail the nature of the bounds we obtain. A measuring device is connected to the physical system to be measured through a transmission line or any kind of transmission medium. The device records the outputs of the system for a short time and reconstructs the state from this data. We assume that the system and the medium operate sufficiently close to an equilibrium point so that a linear model describes them accurately; the models we use are therefore linear port-Hamiltonian systems, see e.g. Cervera et al. (2007), obeying equipartition of energy and fluctuation-dissipation. The transmission medium is

* This work was supported by the Swedish Research Council (VR) under Grant 2007-6350 and Grant 2009-4565, Swedish Foundation for Strategic Research (SSF), and the Belgian Programme on Interuniversity Attraction Poles DYSCO, initiated by the Belgian Federal Science Policy Office. The scientific responsibility rests with its authors. J.-C. D. is a member of naXys (Namur Complex Systems Center).

microscopically modelled by a large number of degrees of freedom which could be, for instance, the movement of the atoms in a steel bar acting as mechanical transmission, or the capacitors and inductors intervening in a usual transmission line, as modelled by the telegrapher's equation. The temperature of the transmission medium imposes a thermal noise to all the degrees of freedom of the medium, which has the effect of disturbing both the measurement device and the measured system. We show that in the absence of information on the initial state of the measured system, the product of measurement accuracy and disturbance on the measured system is in the best case equal to a quantity that is proportional to $2k_B T M^{-1}$, where k_B is Boltzmann's constant, T is the temperature of the medium and M is the inertia of the measured system (typically a mass for a mechanical system). This quantity is independent of the details of the measuring device, the impedance of the transmission medium or the (short) time of measurement. This result is derived in Section 3. An earlier, much less general, version of this result was derived in Sandberg et al. (2011). In particular, here thermal noise is inherent to the transmission medium and is not a result of dissipation in the measurement device. Hence the effect described here applies to all possible measurement devices, and not only to the types discussed in Sandberg et al. (2011). Furthermore, we here also mention how the results generalize to the case when the measured system is not close to equilibrium.

The second main result in this paper takes place when the measured system is known to be a dissipative system at a temperature T' , whose state is confined in a well of energy. In this case, which is typical in many physical situations, the measuring device has a non-zero precision estimate of the initial state to be measured even before starting the measurement, and may exploit this information in order to derive a more favorable trade-off than precedently. Even so, we prove that the ratio between the perturbation of the measured system and the improvement of precision on the estimate is at least $2T/T'$. Thus a large improvement of the variance of the error can only be achieved at the cost of a large perturbation of the system, for given temperatures. This is developed in Section 4.

The effects we describe typically have negligible consequences for large objects, but may be very important for nanoscale machines or biomolecules. We show a numerical example of the measurement of the speed of a blood cell attached to a wall through a macromolecule in Example 3.

2. PROBLEM FORMULATION AND MODELING

We consider systems that can be split up into three different subsystems, see Fig. 1.

A physical system \mathcal{S} which has a property $y(t)$ we would like to measure. We assume the property $y(t)$ is a physical quantity that has a conjugate quantity $u(t)$ such that the pair $u(t)$ and $y(t)$ forms a *port*, see for example Anderson and Vongpanitlerd (2006); Cervera et al. (2007), and the product $y(t)^T u(t)$ is the external work rate on \mathcal{S} .

A measurement device \mathcal{M} which produces an estimate $\hat{y}_m(t)$ of the physical quantity $y(t)$. An optimal measure-

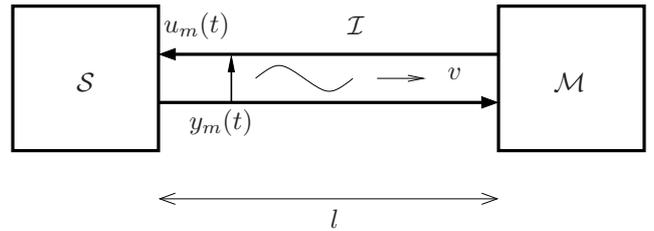


Fig. 1. The measured system \mathcal{S} , the lossless interconnection medium \mathcal{I} , and a measurement device \mathcal{M} . Changes in $y(t)$ and $u(t)$ travel with velocity v through the medium. We only consider times t such that $t < l/v$, such that we do not have to take reflections and initial states in \mathcal{M} into account.

ment device is denoted \mathcal{M}^* , and its estimate is denoted $\hat{y}_m^*(t)$.

A measurement medium \mathcal{I} which interconnects \mathcal{S} and \mathcal{M} . We model the medium using a lossless one-dimensional wave equation, with characteristic impedance $Z > 0$. Perturbations in $u(t)$ and $y(t)$ at the interface to \mathcal{S} travel with a velocity v towards \mathcal{M} . In the absence of reflections and thermal noise, it holds that $u(t) = -Zy(t)$.

Remark 1. We assume $y(t)$ can be chosen as an *output* and $u(t)$ as an *input* of \mathcal{S} . It is clear that this constitutes an restriction and that a behavioral modeling setting, see for example Polderman and Willems (1997), is more general. However, to simplify the presentation we choose this input-output setting. The model of the medium \mathcal{I} is accurate for small perturbations and short time scales where dissipative effects can be neglected, see Remark 3.

Example 1. As an example, $y(t)$ can be the electrical potential difference between two terminals of \mathcal{S} , and $u(t)$ the current flowing into and out of the system. In a mechanical example, $y(t)$ can be the velocity of a surface of \mathcal{S} , and $u(t)$ the corresponding force on the surface. In the electrical case, \mathcal{I} becomes a lossless transmission cable with characteristic impedance Z . In the mechanical case, it becomes a lossless elastic beam with acoustic impedance Z .

Problem formulation: We want to bound the accuracy of the estimate $\hat{y}_m(t)$ for measurements of short duration $t < l/v$, where l is the distance between \mathcal{S} and \mathcal{M} . For such t , the initial state and reflections from \mathcal{M} have no influence on \mathcal{S} . We also want to quantify a limit on the necessary simultaneous perturbation of the physical quantity $y(t)$.

3. DETERMINISTIC PHYSICAL SYSTEM \mathcal{S}

First, we consider deterministic systems \mathcal{S} that can be modeled as linear time-invariant systems,

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), & x(0) &= x_0, \\ y(t) &= Cx(t), & y_0 &:= Cx_0, \end{aligned} \quad (1)$$

for $t \geq 0$, $x(t) \in \mathbb{R}^n$, and $y(t), u(t) \in \mathbb{R}^p$. We assume \mathcal{S} has fixed but unknown (to the measurement device \mathcal{M}) initial state $x_0 \in \mathbb{R}^n$. Let us also define the *inverse inertia* of \mathcal{S} as $M^{-1} := CB$. The solution to (1) for small t and with no input ($u(t) = 0$) is

$$y(t) = Ce^{At}x_0 = y_0 + CAx_0t + O(t^2). \quad (2)$$

This represents the unperturbed behavior of \mathcal{S} ; its behavior had we not tried to measure it using \mathcal{IM} .

Remark 2. If the true system \mathcal{S} is nonlinear, as is frequently the case, a linear approximation in the form (1) is accurate close to an equilibrium point, as is well known. Even if x_0 is far away from an equilibrium point, one can still linearize the nonlinear system \mathcal{S} around a nominal trajectory close to x_0 and obtain a locally accurate linear *time-varying* approximation $\{A(t), B(t), C(t)\}$. If this time-varying realization is smooth and bounded, the results in this section still apply with minor modification using the constant matrices $\{A(0), B(0), C(0)\}$. This is because we only consider short measurement times t . In conclusion, the linearity assumption on \mathcal{S} in (1) is not as serious as it may seem at first.

To estimate the output $y(t)$, we connect \mathcal{S} to a measurement device \mathcal{M}^* using the interconnection medium \mathcal{I} , see Fig. 1. The measurement device \mathcal{M}^* is assumed to have a model of \mathcal{S} and \mathcal{I} , and implements a Kalman filter that gives an optimal estimate of the output. Any other possible measurement device \mathcal{M} that has less information will obtain a worse estimate than \mathcal{M}^* . Hence we can use the idealized measurement device \mathcal{M}^* to quantify the bound on measurement accuracy.

The inherent thermal noise characteristics of \mathcal{I} can then be determined using the well-known *fluctuation-dissipation theorem*, as discussed in Nyquist (1928); Callen and Welton (1951). The terminals of \mathcal{I} then satisfy the following relation.

Proposition 1. Suppose a lossless transmission medium of characteristic impedance $Z = Z^T > 0$, $Z \in \mathbb{R}^{p \times p}$, with respect to the port (u_Z, y_Z) , is in thermal equilibrium with temperature T_Z [Kelvin]. Then it holds

$$y_Z(t) = Zu_Z(t) + \sqrt{2k_B T_Z Z} w_Z(t),$$

where $w(t)$ is unit-intensity white noise ($\mathbf{E}w_Z(t)w_Z(\tau)^T = \delta(t-\tau)I_p$, k_B is Boltzmann's constant, and the unit of the work rate $u_Z(t)^T y_Z(t)$ is Joule/sec.

Remark 3. We choose to work with lossless transmission media, which are characterized by real, frequency-independent characteristic impedances. For instance, a lossless transmission line described by the telegrapher's equation has an impedance $\sqrt{\frac{L}{C}}$, where L is the inductance per unit length and C is the capacitance per unit length. This assumption is not restrictive, at least in first approximation, because we are interested in small measurement times, for which dissipation is negligible. The impedance for a general transmission line according to the telegrapher's equation is $\sqrt{\frac{R+i\omega L}{G+i\omega C}}$, where the resistance per unit length R and the conductance per unit length G express the lossiness of the line, tends to the lossless $\sqrt{\frac{L}{C}}$ for very high frequencies—the only one that matter for a short measurement time. The symmetry of the impedance matrix Z is simply the expression of Onsager's reciprocity relations, see for example Willems (1972).

Next, we use the *physical assumption* that energy is conserved and that the interconnection of \mathcal{S} and \mathcal{I} is modelled as $u(t) = -y_Z(t)$, $u_Z(t) = y(t)$. We obtain the model \mathcal{SI} ,

$$\begin{aligned} \dot{x}_m(t) &= (A - BZC)x_m(t) - B\sqrt{2k_B T_Z Z} w_Z(t), \\ y_m(t) &= Cx_m(t), \quad x_m(0) = x_0, \\ u_m(t) &= -Zy_m(t) - \sqrt{2k_B T_Z Z} w_Z(t), \end{aligned} \quad (3)$$

which is valid as long as no wave from \mathcal{M}^* reaches \mathcal{S} , i.e., time $t < l/v$. The subscript m in $x_m(t)$ etc. is used to distinguish the measured system from the original unperturbed solution $x(t)$, see (1)–(2). If we Taylor-expand the solution to (3), we obtain

$$\begin{aligned} y_m(t) &= Ce^{(A-BZC)t} x_0 \\ &\quad - \int_0^t Ce^{(A-BZC)(t-\tau)} B\sqrt{2k_B T_Z Z} w_Z(\tau) d\tau \\ &= y_0 + C(A - BZC)x_0 t \\ &\quad - CB\sqrt{2k_B T_Z Z} \int_0^t w_Z(\tau) d\tau + O(t\sqrt{t}), \end{aligned}$$

as $t \rightarrow 0$. The conjugate signal $u_m(t)$ is transmitted through the medium \mathcal{I} as a wave, and is received by the measurement device \mathcal{M}^* implementing the Kalman filter at a time instant l/v later. Since \mathcal{M}^* has a perfect model of the system and the medium (apart from its noise realization $w_Z(t)$), the optimal Kalman filter for the estimation of $y_m(t)$ becomes

$$\begin{aligned} \dot{\hat{x}}_m(t) &= (A - BZC)\hat{x}_m(t) + K(t)(u_m(t) - \hat{u}_m(t)), \\ \hat{u}_m(t) &= -ZC\hat{x}_m(t), \\ \hat{y}_m^*(t) &= C\hat{x}_m(t), \end{aligned}$$

where $K(t)$ is the time-varying Kalman gain.

The accuracy of the Kalman filter is given by the covariance matrix $P(t) \in \mathbb{R}^{n \times n}$ that satisfies the differential Riccati equation, see for example Åström (2006),

$$\begin{aligned} \dot{P} &= (A - BZC)P + P(A - BZC)^T \\ &\quad - (PC^T Z - 2k_B T_Z BZ)(2k_B T_Z Z)^{-1} (PC^T Z - 2k_B T_Z BZ)^T \\ &\quad \quad \quad + 2k_B T_Z BZB^T. \end{aligned} \quad (4)$$

We assume that $\mathbf{E}[\hat{y}_m^*(0) - y_0]^2 = \infty$, which means \mathcal{M}^* has no initial knowledge of y_0 and compute the first term in a power-series expansion of the solution $P(t)$ of (4),

$$P(t) = P_{-1}t^{-1} + P_0 + P_1t + O(t^2), \quad t \rightarrow 0.$$

For the first coefficient, which is dominant for short measurements, we obtain the equation

$$P_{-1} = P_{-1}C^T Z C P_{-1} / 2k_B T_Z.$$

Since $\text{Var}[\hat{y}_m^*(t) - y_m(t)] = CP(t)C^T$, the best possible estimation accuracy of the output becomes

$$\text{Var}[\hat{y}_m^*(t) - y_m(t)] = 2k_B T_Z Z^{-1} t^{-1} + O(1). \quad (5)$$

It is seen that a medium of low temperature and large characteristic impedance yields a small lower limit on the accuracy. It is also clear that a longer measurement (larger t), gives a better estimate. Hence, it seems that a large medium impedance Z is preferable. However, as seen next, there are problems with large Z .

Notice that the difference between the unperturbed output $y(t)$ and the expected measured output becomes

$$y(t) - \mathbf{E}y_m(t) = M^{-1}Zy_0t + O(t^2), \quad t \rightarrow 0.$$

We call this quantity the *deterministic back action* of the measurement, which is present even when $T_Z = 0$. When $T_Z > 0$, it creates an uncertainty in y_m , and we obtain

$$\mathbf{E}y_m(t) - y_m(t) = M^{-1}\sqrt{2k_B T_Z Z} \int_0^t w_Z(\tau) d\tau + O(t\sqrt{t}),$$

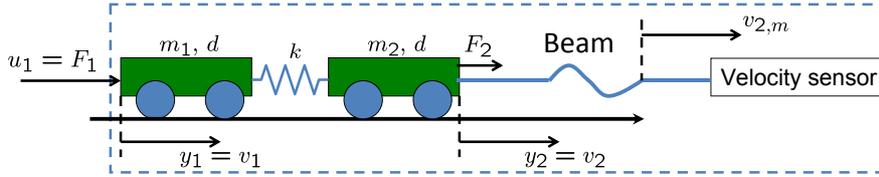


Fig. 2. A cart system. In Example 2, the problem of measuring the second cart's velocity v_2 is considered.

and the variance

$$\text{Var}[y_m(t)] = 2k_B T_Z M^{-1} Z M^{-T} t + O(t^2), \quad (6)$$

is defined as the *stochastic back action* of the measurement. It seen that both the deterministic and stochastic back action are increasing with measurement time t and with impedance Z . Hence, the measurement accuracy (5) is obtained at the expense of perturbation of the system \mathcal{S} , and it is not possible to measure \mathcal{S} without perturbation. We summarize the physical limits we have obtained in the following theorem.

Theorem 2. For the measurement problem defined above, the estimation accuracy of any measurement device \mathcal{M} is bounded as

$$\begin{aligned} \text{Var}[\hat{y}_m(t) - y_m(t)] &\geq \text{Var}[\hat{y}_m^*(t) - y_m(t)] \\ &= 2k_B T_Z Z^{-1} t^{-1} + O(1). \end{aligned}$$

The measurement simultaneously perturbs the output according to

$$\begin{aligned} y(t) - \mathbf{E}y_m(t) &= M^{-1} Z y_0 t + O(t^2), \\ \text{Var}[y_m(t)] &= 2k_B T_Z M^{-1} Z M^{-T} t + O(t^2). \end{aligned}$$

Furthermore, there is a trade-off between measurement accuracy and perturbation (back action),

$$\begin{aligned} \sqrt{\text{Tr Var}[y_m(t)]} \sqrt{\text{Tr Var}[\hat{y}_m(t) - y_m(t)]} \\ \geq 2k_B T_Z \text{Tr } M^{-1} + O(t). \end{aligned} \quad (7)$$

Proof. It remains to prove (7), which is the first main result of this paper. Indeed $2k_B T_Z M^{-1} = UV$, where $U = M^{-1} \sqrt{2k_B T_Z Z} t$ and $V = \sqrt{2k_B T_Z Z^{-1/2} t^{-1/2}}$. From the Cauchy-Schwartz inequality, one gets $2k_B T_Z \text{Tr } M^{-1} \leq \sqrt{\text{Tr } U^T U} \sqrt{\text{Tr } V^T V}$, which is, up to $O(t)$, equal to $\sqrt{\text{Tr Var}[y_m(t)]} \sqrt{\text{Tr Var}[\hat{y}_m^*(t) - y_m(t)]}$ which is a lower bound on $\sqrt{\text{Tr Var}[y_m(t)]} \sqrt{\text{Tr Var}[\hat{y}_m(t) - y_m(t)]}$.

The observer effect is therefore dependent on both the temperature T_Z of the measurement medium and the 'size' M of the measured system. We consider the following numerical example to evaluate its importance in a practical situation.

Example 2. Consider the double cart system in Fig. 2. Assume we would like to measure the velocity $y = v_2$ of the second cart. The interconnection to the velocity sensor is modelled by a lossless elastic beam with acoustic impedance Z . The carts can be modelled using the matrices

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -k/m_1 & -d_1/m_1 & k/m_1 & 0 \\ 0 & 0 & 0 & 1 \\ k/m_2 & 0 & -k/m_2 & -d_2/m_2 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1/m_2 \end{pmatrix}$$

$$C = (0 \ 0 \ 0 \ 1),$$

where m_1, m_2 denote the masses of the carts, k the spring constant, and d_1, d_2 viscous friction. The conjugate

variable to v_2 is the force on the second cart, $u = F_2$. Using this input and output, we obtain the inertia $M = m_2$, i.e., the mass of the second cart. Assume now the second cart initially has a positive velocity, $y_0 = v_{2,0} > 0$, and interconnection to the measurement medium and device occurs at $t = 0$. The bounds in Theorem 2 now say the expected decrease of velocity after a short measurement over $[0, t]$ is $Z v_{2,0} t / m_2$, and the variance in the cart velocity is $2k_B T_Z Z t / m_2^2$. At the same time, no velocity sensor can obtain a velocity estimate with a variance smaller than $2k_B T_Z / Z t$. The product of the corresponding standard deviations is $2k_B T_Z / m_2$, which is $8.2810^{-21} \text{ m}^2/\text{s}^2$ for $T_Z = 300 \text{ K}$ and $m_2 = 1 \text{ kg}$. It is therefore negligible. However, for a mass of 10^{-23} kg , which is realistic for a nanomachine, this product becomes $828 \text{ m}^2/\text{s}^2$, which is not negligible any more.

Remark 4. We have considered above that the measurement time is so short that there is no back action of the measurement device \mathcal{M} onto the system \mathcal{S} . Due to the nonzero length l of the transmission medium \mathcal{I} , the information takes a certain time $\delta = l/v$ to travel from the measured system to the measuring device; thus the estimate of the state at time t is computed by the measuring device only at time $t + \delta$. If the measuring time is less than δ , then no information has time to travel from the measurement device to the measured system; if we suppose a measuring time larger than δ , of course this would only increase the back action and make the trade-off even worse.

4. PHYSICAL SYSTEM \mathcal{S} IN THERMAL EQUILIBRIUM

In the previous section, we assumed the state of the system \mathcal{S} was completely unknown to the measurement device. Many times this is not a realistic assumption. For example, in many cases it is reasonable to assume that the system \mathcal{S} is in a thermal equilibrium with its environment, and its state stays close to a fixed point of the dynamics, in other words remain confined in a well of potential energy. If the temperature is known, the measurement device can use this information. But as we see next, there are still very similar trade-offs between estimation accuracy and back action.

As argued in, for example, Cervera et al. (2007), a natural framework for modeling of physical systems is a port-Hamiltonian model framework, which is a generalization of the Hamiltonian systems often encountered in physics. This is also a suitable framework for studying systems in thermal equilibrium. A linear port-Hamiltonian system without any dissipation takes the form

$$\begin{aligned} \dot{x}(t) &= JQx(t) + Bu(t) \\ y(t) &= B^T Qx(t), \end{aligned} \quad (8)$$

where J is anti-symmetric $J = -J^T$, and $H(x) = \frac{1}{2}x^T Qx$ is the Hamiltonian of the system, where Q is assumed to be symmetric positive definite, $Q = Q^T > 0$. The inverse inertia is here defined as $M^{-1} := B^T Q B$. Indeed typically $B^T Q B$ has the physical interpretation of a partial inverse of the inertia matrix for port-Hamiltonian systems¹. For example, in a typical mechanical system context, $Q = \text{diag}\{K_i, M_i^{-1}\}$ where K_i is the stiffness matrix (springs) and M_i is the inertia (mass) matrix, and then $B^T Q B$ contains elements of M_i^{-1} . The pair (y, u) constitutes a port, and since the system has no dissipation, it holds that

$$\frac{dH(x(t))}{dt} = y(t)^T u(t),$$

and the amount of energy stored is given by H . Next, we split the port (y, u) into two pairs (y_K, u_K) and (y_Z, u_Z) , using the partition $B = [B_K \ B_Z]$.

Remark 5. That we here consider linear port-Hamiltonian systems, and not more general nonlinear ones, is justified since we next assume the systems are in thermal equilibrium and hence are close to their equilibrium points.

Assume now the port variables (y_K, u_K) in the system (8) are connected to a dissipative element of temperature T_K that satisfies the fluctuation-dissipation theorem (Proposition 1), $u_K(t) = -K y_K(t) - \sqrt{2k_B T_K} \bar{K} w_K(t)$. This results in a generalized Langevin equation,

$$\dot{x}(t) = (J - B_K K B_K^T) Q x(t) + B_Z u_Z(t) - B_K \sqrt{2k_B T_K} \bar{K} w_K(t), \quad (9)$$

$$y_Z(t) = B_Z^T Q x(t).$$

In the following, we will model the unmeasured system \mathcal{S} using (9) and putting $u_Z = 0$. The following lemma guarantees that the system \mathcal{S} is stabilized by the dissipative element if the realization is controllable.

Lemma 3. If (JQ, B_K) is controllable and $K = K^T > 0$, then $(J - B_K K B_K^T)Q$ is Hurwitz.

In steady state, the variance of the state vector in the unmeasured \mathcal{S} , $P_\infty = \lim_{t \rightarrow \infty} \mathbf{E}x(t)x(t)^T$, satisfies the Lyapunov equation

$$(J - B_K K B_K^T) Q P_\infty + P_\infty Q (J - B_K K B_K^T)^T + 2k_B T_K B_K K B_K^T = 0,$$

with the unique solution (by Lemma 3) $P_\infty = k_B T_K Q^{-1}$. In fact, we *define* the temperature of the port-Hamiltonian system using this observation.

Definition 1. A port-Hamiltonian system \mathcal{S} is of temperature T if its steady-state state covariance is in the form $P_\infty = k_B T Q^{-1}$. \mathcal{S} is then said to be in *thermal equilibrium*.

This definition conforms well with physical insight. For example, the expected value of the (quadratic) Hamiltonian becomes

$$\mathbf{E}H(t) = \frac{1}{2} \text{Tr}[Q P_\infty] = \frac{n}{2} k_B T_K, \quad (10)$$

which is what the *equipartition theorem* found in the physics literature also predicts.

If the system \mathcal{S} in (9) is connected to a measurement medium \mathcal{I} using the port variables (y_Z, u_Z) , we obtain the model

$$\begin{aligned} \dot{x}_m(t) &= (J - B_K K B_K^T - B_Z Z B_Z^T) Q x_m(t) \\ &\quad - B_K \sqrt{2k_B T_K} \bar{K} w_K(t) - B_Z \sqrt{2k_B T_Z} \bar{Z} w_Z(t), \\ y_m(t) &= B_Z^T Q x_m(t), \quad x_m(0) = x_0, \\ u_m(t) &= -Z y_m(t) - \sqrt{2k_B T_Z} \bar{Z} w_Z(t). \end{aligned} \quad (11)$$

This can be compared to the model (3). Assume now the system is in steady-state of temperature T_K before the measurement starts at time $t = 0$. The question is if we still obtain trade-offs similar to those in Section 3.

An expansion of the solutions to (9) (putting $u_Z = 0$) and (11) yields

$$\begin{aligned} (y - y_m)(t) &= B_Z^T Q B_Z \sqrt{2k_B T_Z} \bar{Z} \int_0^t w_Z(\tau) d\tau \\ &\quad + B_Z^T Q B_Z Z B_Z^T Q x_0 t + O(t\sqrt{t}), \end{aligned}$$

assuming $w_K(t)$ are identical. So the back action, the difference between unmeasured and measured \mathcal{S} , becomes

$$\begin{aligned} \mathbf{E}[(y - y_m)(t)] &= M_Z^{-1} Z y_0 t + O(t^2), \\ \text{Var}[(y - y_m)(t)] &= 2k_B T_Z M_Z^{-1} Z M_Z^{-1} t + O(t^2), \end{aligned}$$

where $M_Z^{-1} = B_Z^T Q B_Z$ is the inverse inertia of \mathcal{S} , as seen from the measurement port variables (y_Z, u_Z) .

A Taylor expansion of the solution to the Kalman filter differential Riccati equation for (11),

$$P(t) = P(0) + P_1 t + O(t^2) = k_B T_K Q^{-1} + P_1 t + O(t^2),$$

yields that

$$P_1 = -\frac{1}{2} \frac{k_B T_K^2}{T_Z} B_Z Z B_Z^T. \quad (12)$$

Here we have assumed \mathcal{S} has temperature T_K at time $t = 0$ ($P(0) = k_B T_K Q^{-1}$) before the measurement starts. Thus the uncertainty in the estimation of the output becomes

$$\begin{aligned} \text{Var}[(\hat{y}_m^* - y_m)(t)] &= B_Z^T Q P(t) Q B_Z \\ &= k_B T_K M_Z^{-1} - \frac{1}{2} \frac{k_B T_K^2}{T_Z} M_Z^{-1} Z M_Z^{-1} t + O(t^2). \end{aligned}$$

This relation quantifies how much the estimation uncertainty *decreases* after a measurement of length t using an optimal measurement device. We summarize the physical limits for port-Hamiltonian systems in thermal equilibrium in the following theorem.

Theorem 4. Suppose the port-Hamiltonian system \mathcal{S} has temperature T_K [Kelvin] at time $t = 0$. For the measurement problem defined above, the improvement of estimation accuracy for any measurement device \mathcal{M} ,

$$\Delta \text{Var}(t) := k_B T_K M_Z^{-1} - \text{Var}[(\hat{y}_m - y_m)(t)],$$

is bounded as

$$\begin{aligned} \Delta \text{Var}(t) &\leq \Delta \text{Var}^*(t) = k_B T_K M_Z^{-1} - \text{Var}[(\hat{y}_m^* - y_m)(t)] \\ &= \frac{1}{2} \frac{k_B T_K^2}{T_Z} M_Z^{-1} Z M_Z^{-1} t + O(t^2). \end{aligned} \quad (13)$$

The measurement simultaneously perturbs the output according to

$$\begin{aligned} \mathbf{E}[(y - y_m)(t)] &= M_Z^{-1} Z y_0 t + O(t^2), \\ \text{Var}[(y - y_m)(t)] &= 2k_B T_Z M_Z^{-1} Z M_Z^{-1} t + O(t^2). \end{aligned} \quad (14)$$

Hence, there is a price to pay for measurement accuracy in terms of perturbation (back action), and it holds

¹ The authors would like to thank an anonymous reviewer for pointing this out.

$$\begin{aligned} \sqrt{\text{Tr Var}[(y - y_m)(t)]} &= \frac{2T_Z}{T_K} \sqrt{\text{Tr } \Delta \text{Var}^*(t)} + O(t\sqrt{t}) \\ &\geq \frac{2T_Z}{T_K} \sqrt{\text{Tr } \Delta \text{Var}(t)} + O(t\sqrt{t}). \end{aligned} \quad (15)$$

This is the second main result of the paper. Remark that (15) is consistent with (7). Indeed, as T_K grows to infinity, the initial variance of the state grows to infinity. The improvement of variance grows to infinity as well, and the ratio above tends towards zero. Here however, the size of the measured system does not play a role anymore. When the temperature T_K of the measured system is small, the trade-off becomes unfavorable: this comes from the fact that the initial state is then already well known with a small variance of the error; therefore only a small improvement of the error variance is to be expected.

Example 3. We consider a system which is inspired from Kamiti and van de Ven (1996). There, a blood cell is connected to a wall with a macromolecular spring. We model the blood cell as a point mass m connected through a linear spring constant k . The cell is immersed in a fluid with viscous friction $K = f$ and temperature T_K . The system can be modelled as a port-Hamiltonian system where

$$\begin{aligned} x &= \begin{pmatrix} x_1 \\ p_1 \end{pmatrix}, \quad J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \\ Q &= \begin{pmatrix} k & 0 \\ 0 & 1/m \end{pmatrix}, \quad B_K = (0 \ 1)^T, \\ u_K &= F, \quad y_K = B_K^T Q x = p_1/m = v, \end{aligned}$$

where F is the frictional force on the blood cell, and v its velocity. For viscous friction with thermal noise, we obtain

$$u_K = -f p_1/m - \sqrt{2k_B T_K} f w_K.$$

At time $t = 0$, we assume the blood cell is brought into contact with a mechanical measurement probe with acoustic impedance Z . The measurement port and inverse inertia then becomes

$$B_Z = (0 \ 1)^T, \quad M_Z^{-1} = B_Z^T Q B_Z = 1/m.$$

The blood cell has the mass $m = M_Z = 2.3 \cdot 10^{-13}$ kg, and the molecular spring has $k = 1.9 \cdot 10^{-6}$ N/m. Let us assume the environment has room temperature $T_K = 300$ K. The fluid has viscous friction $f = 1.63 \cdot 10^{-7}$ Ns/m, see Kamiti and van de Ven (1996). In thermal equilibrium, this gives

$$\mathbf{E}v(0) = 0 \text{ m/s} \quad \mathbf{E}v(0)^2 = 1.78 \cdot 10^{-8} \text{ m}^2/\text{s}^2.$$

Next, we compute the exact back action and estimation uncertainties by numerically solving the corresponding differential Lyapunov and Riccati equations, and compare with the Taylor approximations in Theorem 4. Let us first assume the measurement probe has the same temperature and impedance as the fluid, i.e., $T_Z = T_K$ and $Z = f$. The blood cell experiences this as a doubling of the friction when the measurement starts. In Figs. 3–4, the estimation accuracy and the back action are shown as functions of measurement time t . For small times t , in this case about $t < 0.2 \mu\text{s}$, the curves are close to being affine and the approximation predicts the bounds well. Changes in Z and T_Z will change the slopes of the curves, as predicted in Theorem 4. That is, increased accuracy in Fig. 3 is only obtained at the expense of larger perturbation in Fig. 4.

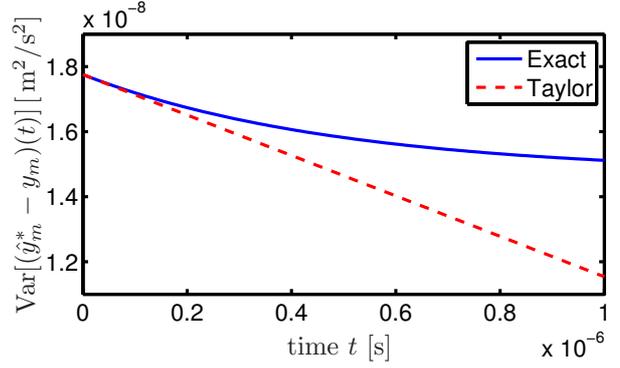


Fig. 3. The variance of the estimation error of an optimal velocity measurement device. Exact and affine approximation shown, see (13).

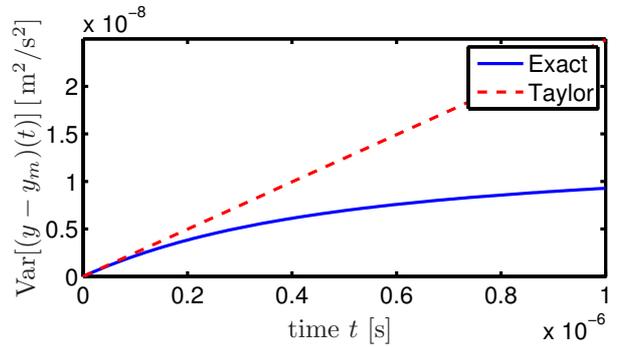


Fig. 4. The stochastic back action. Exact and linear approximation shown, see (14).

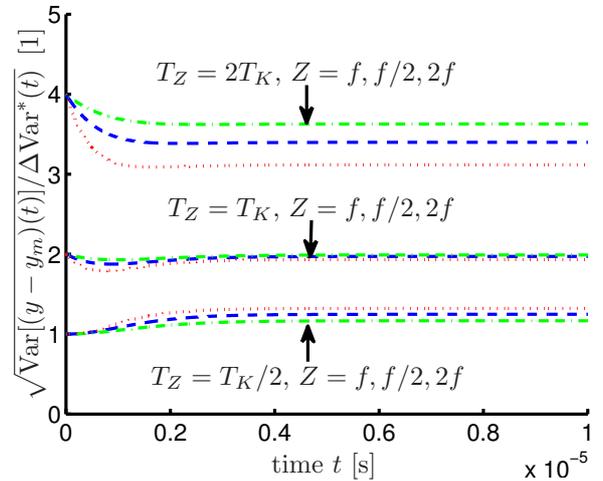


Fig. 5. The ratio between stochastic back action and estimation uncertainty improvement is approximately constant and equal to $2T_Z/T_K$, see (15), far beyond the affine regime $t < 0.2 \mu\text{s}$ found in Figs. 3–4 (note maximum time is here $10 \mu\text{s}$).

Interestingly, the trade-off (15) seems to extend rather well far beyond the affine regime $t < 0.2 \mu\text{s}$, see Fig. 5. This is especially true when T_Z and T_K are close, as seen in the figure. Furthermore, it is seen that the influence of Z on the trade-off is small (the acoustic impedance of the probe), just as predicted in (15).

Remark 6. It is tempting to try to avoid back action by using a more noninvasive measuring device than a probe. For example, one could put a microphone in the fluid to record the pressure waves created by the cell. Note, however, that this only moves the interface between the system \mathcal{S} and medium \mathcal{I} away from the blood cell. The system \mathcal{S} would now include the blood cell and the surrounding fluid which together also can be modelled by a much larger port-Hamiltonian system. The membrane of the microphone becomes \mathcal{I} and has its own acoustic impedance and temperature. So the previous results hold and bound the accuracy with which we can estimate the pressure wave. Based on this estimate, one could use the model \mathcal{S} to estimate the velocity of the blood cell. An optical measuring device could be modelled similarly, but would require that we model interaction with electromagnetic fields as well.

5. CONCLUSIONS

We have exhibited a general form of the observer effect, so general indeed that its conclusion seems unescapable in any practical situation. As long as the measuring device and the measured system are physically separated, they must be interconnected through a transmission medium. As the media considered here are of linear nature, it seems tempting to use nonlinear, e.g., digital channels instead. But then an analog transmission line is likely to exist in the analog/digital converter. Although our analysis is confined to the case of short measurements, numerical experimentations in the last section suggest that the main conclusion remains true for long time measurements. Future work will be devoted to finding the weakest assumptions under which the observer effect holds.

ACKNOWLEDGEMENTS

Thanks to John C. Doyle for many fruitful and inspiring discussions.

REFERENCES

- Anderson, B.D.O. and Vongpanitlerd, S. (2006). *Network Analysis and Synthesis: A Modern Systems Theory Approach*. Dover Publications.
- Åström, K.J. (2006). *Introduction to Stochastic Control Theory*. Dover Publications.
- Barnes, R.B. and Silverman, S. (1934). Brownian motion as a natural limit to all measuring processes. *Reviews of Modern Physics*, 6, 162–193.
- Callen, H.B. and Welton, T.A. (1951). Irreversibility and generalized noise. *Physical Review*, 83(1), 34–40.
- Cervera, J., van der Schaft, A.J., and Baños, A. (2007). Interconnection of port-Hamiltonian systems and composition of Dirac structures. *Automatica*, 43, 212–225.
- Kamiti, M. and van de Ven, T.G. (1996). Measurement of spring constants of polyacrylamide chains bridging particles to a solid surface. *Macromolecules*, 29, 1191–1194.
- Martins, N.C., Dahleh, M.A., and Doyle, J.C. (2007). Fundamental limitations of disturbance attenuation in the presence of side information. *IEEE Transactions on Automatic Control*, 52(1), 56–66.
- McCombie, C. (1953). Fluctuation theory in physical measurements. *Reports on Progress in Physics*, 16, 266–320.
- Nair, G.N., Evans, R.J., Mareels, I.M.Y., and Moran, W. (2004). Topological feedback entropy and nonlinear stabilization. *IEEE Transactions on Automatic Control (special issue on Networked Control Systems)*, 49(9), 1585–1597.
- Nyquist, H. (1928). Thermal agitation of electrical charge in conductors. *Physical Review*, 32, 110–113.
- Polderman, J.W. and Willems, J.C. (1997). *Introduction to Mathematical Systems Theory — A Behavioral Approach*. Springer.
- Sahai, A. and Mitter, S. (2006). The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link Part I: scalar systems. *IEEE Transactions of Information Theory*, 52, 3369–3395.
- Sandberg, H., Delvenne, J., and Doyle, J.C. (2011). On lossless approximations, the fluctuation-dissipation theorem, and limitations of measurements. *IEEE Transactions on Automatic Control*, 56(2), 293–308. doi: 10.1109/TAC.2010.2056450.
- Sinopoli, B., Schenato, L., Franceschetti, M., Poolla, K., Jordan, M.I., and Sastry, S.S. (2003). Kalman filtering with intermittent observations. *IEEE Transactions Automatic Control*, 49(9), 1453–1464.
- Willems, J.C. (1972). Dissipative dynamical systems part II: Linear systems with quadratic supply rates. *Archive for Rational Mechanics and Analysis*, 45, 352–393.