Nonlinear Feedback Systems Perturbed by Noise: Steady-State Probability Distributions and Optimal Control

Daniel Liberzon, *Member, IEEE* and Roger W. Brockett, *Fellow, IEEE*

***Abstract*—In this paper we describe a class of nonlinear feedback systems perturbed by white noise for which explicit formulas for steady-state probability densities can be found. We show that this class includes what has been called monotemperaturic systems in earlier work and establish relationships with Lyapunov functions for the corresponding deterministic systems. We also treat a number of stochastic optimal control problems in the case of quantized feedback, with performance criteria formulated in terms of the steady-state probability density.**

***Index Terms*—Nonlinear feedback system, quantizer, steady-state probability density, stochastic optimal control, white noise.**

I. Introduction

The study of linear systems excited by white noise is greatly facilitated by the fact that one can explicitly solve the Fokker-Planck equation which describes the time evolution of the probability density. For nonlinear systems the situation is quite different: not only the transient solutions but even the steady-state ones are difficult to find. Various methods have been used to prove the existence of steady-state probability distributions for quite general classes of nonlinear systems, but such results do not provide specific expressions for steady-state probability densities (this work can be traced back to the 1960's—see the references in [18] and [26]). On the other hand, explicit formulas for steady-state densities have been obtained for certain special classes of systems, most notably gradient systems and Hamiltonian systems with certain types of dissipation [12], [14], [16], [27]. However, the gap between the category of systems for which steady-state densities are known to exist and that of systems for which specific formulas are available is still quite large, which justifies further investigation of the problem.

In this paper we study Itô stochastic systems of the form

$$dx = Ax\,dt + G\,dw + bf(c^T x)\,dt \tag{1}$$

Manuscript received March 20, 1998; revised April 30, 1999. Recommended by Associate Editor, Q. Zhang. The work of D. Liberzon was supported in part by ARO DAAH04-95-1-0114, NSF ECS 9634146, and AFOSR F49620-97-1-0108. The work of R. W. Brockett was supported in part by the U. S. Army under Contracts DAAH 04-96-1-0445, DAAG 55-97-1-0114, and DAAL 03-92-G-0115.

D. Liberzon is with the Department of Electrical Engineering, Yale University, New Haven, CT 06520-8267 USA (e-mail: daniel.liberzon@yale.edu).

R. W. Brockett is with the Division of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138 USA (e-mail: brockett@hrl.harvard.edu).

Publisher Item Identifier S 0018-9286(00)05964-X.

where $x, b, c \in \mathbb{R}^n$, w is a standard m-dimensional Wiener process, and A and G are matrices of appropriate dimensions. Such systems arise naturally in control theory; they correspond to feedback systems of Lur'e type perturbed by white noise. Note that the system (1) is not restricted to be of a gradient type, so the available results on steady-state densities for gradient systems are in general not applicable. Our main goal is to address the following three questions.

1. Under what conditions can the steady-state probability density for the system (1) be computed explicitly?
2. Do there exist physically and mathematically meaningful interpretations of these conditions?
3. Provided that these conditions are satisfied, how do the steady-state properties change as the parameters of the system vary?

In the rest of this section we give a more precise formulation of the problem and a detailed outline of the paper. Some of the results reported here have been announced previously in [8] and [9].

The *Fokker–Planck operator* associated with (1) is given by the formula

$$\begin{aligned} L\rho = &-\left(\text{tr}\,A + \sum_{i=1}^{n} b_i c_i f'(c^T x)\right)\rho \\ &- \left(\sum_{i,j=1}^{n} A_{ij} x_j \rho_{x_i} + \sum_{i=1}^{n} b_i f(c^T x)\rho_{x_i}\right) \\ &+ \frac{1}{2}\sum_{j,k=1}^{n} (GG^T)_{jk}\rho_{x_j x_k}. \end{aligned} \tag{2}$$

The problem under consideration is that of finding an explicit formula for a steady-state probability density associated with (1). This amounts to solving the steady-state version of the Fokker-Planck equation

$$L\rho(x) = 0 \tag{3}$$

subject to the constraints $\rho(x) > 0\ \forall\, x$ and $\int_{\mathbb{R}^n} \rho(x)\,dx = 1$. It is reasonable to assume that the function f is sufficiently regular (e.g., continuously differentiable) so that the solutions of (1) and the Fokker-Planck operator (2) are well defined. However, if the results to follow are to be interpreted in a formal sense, this requirement may be relaxed (cf. Section II). Unless explicitly stated otherwise, we make the following two assumptions.

a) All eigenvalues of A have negative real parts.
b) (A, G) is a controllable pair.

Let us first recall what happens in the linear case ($b = 0$). As is well known, under the above assumptions the (unique) steady-state probability density is a Gaussian

$$\rho(x) = \frac{1}{\sqrt{(2\pi)^n \det Q}} e^{-(1/2)x^T Q^{-1} x}$$

where Q is the positive definite symmetric steady-state variance matrix satisfying the equation

$$AQ + QA^T + GG^T = 0. \quad (4)$$

In this paper we will be concerned with extending this result to nonlinear feedback systems of the form (1). In Section II we study the case when the nonlinearity is given by a piecewise constant function of a certain type, called a *quantizer*. We formulate a condition on the parameters of the system which enables us to obtain an explicit formula for a steady-state probability density. This condition will be called the *compatibility condition*, and systems for which it is satisfied will be called *compatible*. Steady-state densities for compatible quantized feedback systems turn out to be piecewise Gaussian. In Section III we show that the same condition leads to an explicit formula for a steady-state density in the general nonlinear feedback case, thereby answering question 1 posed above. The steady-state densities that we obtain can be viewed as being of the Maxwell-Boltzmann type. Some examples are given in Section IV.

We then turn to question 2. A physical interpretation of compatibility is given in Section V in terms of certain concepts from statistical thermodynamics. Namely, systems that are monotemperaturic in the sense of [11] turn out to be compatible. This shows more clearly the place that compatible systems occupy among all systems of the form (1). In Section VI we single out a class of systems for which the compatibility condition takes a particularly transparent form and show how the steady-state probability densities are related to Lyapunov functions for deterministic nonlinear feedback systems. This makes contact with the Lur'e problem of absolute stability.

Some extensions and related issues are discussed next. In Section VII we address the question of convergence of the probability density associated with (1) to steady state and give a brief overview of available results on existence and uniqueness of steady-state probability distributions. In Section VIII we consider systems with unstable linear part. In Section IX we use the work of Zakai on the existence of steady-state probability distributions to obtain bounds on second moments for certain noncompatible systems. This provides a natural generalization of the results of Section V to a class of systems whose temperature is not constant, but rather varies along a certain sufficiently small interval.

The last two sections are devoted to question 3. Namely, we study how the steady-state behavior in the case of quantized feedback depends on the parameters of a given compatible system. In Section X we consider several performance criteria of the quadratic-gaussian type. Finally, in Section XI we consider another performance criterion, related to the number of switching hyperplane crossings per unit time. This leads to an interesting optimal control problem, which can be interpreted as minimization of the cost of implementing a feedback control law, and involves a novel application of Rice's formula. Throughout the last five sections of the paper, some open problems for future work are also pointed out.

II. Quantized Feedback Systems

Let us denote by Q the solution of (4). Of course, it is not the steady-state variance matrix anymore in the nonlinear case. We will say that the system (1) is *compatible* if the following *compatibility condition* is satisfied:

$$b = \lambda AQc \quad \text{for some } \lambda \in \mathbb{R}. \quad (5)$$

We will be particularly interested in the case when the nonlinearity is given by a piecewise constant function defined as follows. Given a positive integer M and a nonnegative real number Δ, we define the *quantizer* q: $\mathbb{R} \to \mathbb{Z}$ with *sensitivity* Δ and *saturation value* M by the formula

$$q(z) = \begin{cases} M, & \text{if } z \geq (M+1/2)\Delta \\ -M, & \text{if } z < -(M+1/2)\Delta \\ \left\lfloor \dfrac{z}{\Delta} + \dfrac{1}{2} \right\rfloor, & \text{if } -(M+1/2)\Delta \leq z < (M+1/2)\Delta \end{cases}$$

In other words, on the interval $J_k := [(k-1/2)\Delta, (k+1/2)\Delta)$, where $k \in \mathbb{Z}$ and $-M \leq k \leq M$, the quantizer q takes on the value k, and for $|z| > (M+1/2)\Delta$ the quantizer saturates. More generally, suppose that we have n quantizers q_i: $\mathbb{R} \to \mathbb{Z}$ with sensitivities Δ_i and saturation values M_i, $i = 1, \cdots, n$. We define a quantizer q: $\mathbb{R}^n \to \mathbb{Z}^n$ by $q(x) := (q_1(x_1), \cdots, q_n(x_n))$, where $(x_1, \cdots, x_n)$ are the coordinates of x relative to a fixed orthonormal basis in $\mathbb{R}^n$. If all q_i's have the same sensitivity Δ, we will call q a *uniform* quantizer with sensitivity Δ. The above notation is similar to the one used by Delchamps in [15].

Consider the system

$$dx = Ax\, dt + G\, dw + bq(c^T x)\, dt \quad (6)$$

with q being a quantizer with sensitivity $\Delta > 0$ and saturation value M. We will sometimes allow the possibility that q has an infinite set of values, i.e., $M = \infty$, in which case we will further assume that λ in (5) is positive. We will not embark on the issue of existence of solutions for stochastic differential equations with discontinuous right-hand side such as (6). The situation when instead of a piecewise constant function q one uses a suitable smooth approximation is covered by the existing theory. The steady-state probability density associated with (6) is to be understood as a solution of the equation (3) almost everywhere and can be obtained in the limit as smooth approximations approach q. The problem of obtaining solutions of the Fokker-Planck equation for the system (6) makes contact with the work reported in [20]; see [10] and the references therein for a discussion of quantized feedback systems and their importance in applications.

It is not hard to show that the function defined by

$$\rho(x) = N \exp[-(1/2)\,(x + A^{-1}bk)^T Q^{-1}(x + A^{-1}bk) + d_k] \quad \text{if } c^T x \in J_k \quad (7)$$

with arbitrary constants N and d_k satisfies the equation for a steady-state probability density associated with (6) almost everywhere. This function is piecewise Gaussian. Clearly, if $x \in \mathbb{R}$, we can always determine particular values of d_k so as to make ρ continuous. However, this is not necessarily true in the multidimensional case. As we now show, the compatibility condition (5) is precisely what makes it possible to obtain a continuous steady-state probability density.

Theorem 1: If the compatibility condition (5) is satisfied, then the process described by the system (6) admits a steady-state probability distribution with a continuous piecewise Gaussian density.

Proof: Define

$$\rho(x) = N \exp\left[-\tfrac{1}{2}x^T Q^{-1} x - \lambda \int_0^{c^T x} q(v)\, dv\right] \tag{8}$$

where $N > 0$ is fixed by the requirement that $\int_{\mathbb{R}^n} \rho(x)\, dx = 1$. Such a normalization constant N always exists. Indeed, we have assumed that either $M < \infty$, or $M = \infty$ and then $\lambda > 0$. In each of these cases the function defined by the exponential in (8) belongs to $L^1(\mathbb{R}^n)$. In light of (5), the function (8) is a special case of (7) with $d_k = (k^2/2)(b^T(A^{-1})^T Q^{-1} A^{-1} b + \lambda\Delta)$. The statement of the theorem follows. □

The above result actually holds in the case of an arbitrary piecewise constant function q. In the next section we will see how it can be extended to the general nonlinear feedback system (1).

Now suppose that we are given $m \le n$ quantizers $q_1, \cdots, q_m$ and m linearly independent vectors $c_1, \cdots, c_m \in \mathbb{R}^n$. The statement of Theorem 1 can be generalized as follows: *the process described by the system*

$$dx = Ax\, dt + G\, dw + \sum_{i=1}^{m} b_i q_i(c_i^T x)\, dt \tag{9}$$

admits a piecewise Gaussian steady-state probability density if $b_i = \lambda_i AQc_i$, $\lambda_i \in \mathbb{R}$ $(i = 1, \cdots, m)$. In this case, the steady-state density takes the form

$$\rho(x) = N \exp\left[-\frac{1}{2}x^T Q^{-1} x - \sum_{i=1}^{m} \lambda_i \int_0^{c_i^T x} q_i(v)\, dv\right].$$

Thus for the quantized state feedback system

$$dx = Ax\, dt + G\, dw + Bq(x)\, dt \tag{10}$$

where q is a uniform quantizer, we need

$$B^{(i)} = \lambda_i (AQ)^{(i)} \quad \text{for some } \lambda_i \in \mathbb{R}\ (i = 1, \cdots, n) \tag{11}$$

where the superscript (i) denotes the ith column of a matrix. The formula (11) can be rewritten as

$$B = AQD, \qquad D = \mathrm{diag}(\lambda_1, \cdots, \lambda_n). \tag{12}$$

III. General Nonlinear Feedback Systems

Consider the system (1) and assume that the compatibility condition (5) holds. Let us look for a steady-state probability density taking the form

$$\rho(x) = N \exp[-\tfrac{1}{2}x^T Q^{-1} x - \lambda F(c^T x)] \tag{13}$$

where $F(z) := \int_0^z f(v) dv$ and $N > 0$ is a normalization constant. Note that (13) includes (8) as a special case. We need to assume that the function defined by the above exponential belongs to $L^1(\mathbb{R}^n)$. This will always be true if, for example, $\lambda > 0$ and $xf(x) \ge 0$ for all x.

Theorem 2: If the compatibility condition (5) is satisfied, then the function ρ given by (13) is a steady-state probability density for the process described by (1).

Proof: Let us evaluate the expression (2) for the Fokker-Planck operator associated with (1) when ρ is given by (13). Combining the terms and making use of (5), we obtain

$$\begin{aligned} L\rho = \rho[(&-\mathrm{tr}\, A - \tfrac{1}{2}\, \mathrm{tr}\, GG^T Q^{-1}) + (\lambda^2 (f(c^T x))^2 \\ &- \lambda f'(c^T x))c^T(\tfrac{1}{2}\, GG^T + AQ)c \\ &+ x^T(Q^{-1}A + \tfrac{1}{2}\, Q^{-1}GG^T Q^{-1})x \\ &+ \lambda f(c^T x)c^T(A + GG^T Q^{-1} + QA^T Q^{-1})x]. \end{aligned}$$

Now it is easy to see that all the terms equal zero because of (4). □

All the exact steady-state solutions obtained to date that we are aware of are closely related to the Maxwell-Boltzmann distribution (see, e.g., [16]). Not surprisingly, so is the steady-state density (13). It would be interesting to have a complete picture of how systems that are compatible in the sense defined here are related to systems for which explicit steady-state probability densities have been obtained previously, e.g., Hamiltonian systems with dissipation studied in [27]. We will make some remarks on this in Section VI. The usefulness of our result stems from the fact that it applies to systems of the form (1) which is natural from the control-theoretic point of view and is not tailored to any special coordinates in which the system takes some canonical form.

Remark 1: Suppose that we write $f(c^T x) = kc^T x + g(c^T x)$, for some number k and a suitable function g, and rewrite the system (1) accordingly as

$$dx = (A + kbc^T)x\, dt + G\, dw + bg(c^T x). \tag{14}$$

Assume that the matrix $A + kbc^T$ is stable (this will certainly be true if k is sufficiently small). As is straightforward to verify, if the original system (1) is compatible, then so is (14). The quadratic term in the expression for the steady-state density associated with (14) will be $-\frac{1}{2}\, x^T(Q^{-1} + \lambda kcc^T)x$. We can thus say that compatibility is preserved under linear feedback transformations. This important property will be implicitly used several times in the sequel.

We can switch to new coordinates in which $Q = TI$ for some $T > 0$. The structure of compatible systems is then revealed by the following statement.

Corollary 3: The system

$$\begin{aligned} dx = &\left(\Omega - \frac{1}{2T}\, GG^T\right)\, x\, dt + G\, dw \\ &+ \lambda T\left(\Omega - \frac{1}{2T}\, GG^T\right)\, cf(c^T x)\, dt \end{aligned} \tag{15}$$

where $\Omega = -\Omega^T$, is compatible.

IV. Examples

It is insightful to see how known solutions to certain problems are captured as special cases of Theorem 2.

Example 1: Consider the system

$$dx_i = f_i(x_1, \cdots, x_n)\, dt + g\, dw_i, \qquad i = 1, \cdots, n \tag{16}$$

where g is a constant and w_i's are independent scalar Wiener processes. We will call such a system *gradient* if there exists a function $\phi(x_1, \cdots, x_n)$ such that $f_i = -(\partial\phi/\partial x_i)$. It is well known and straightforward to verify that a steady-state density is then given by the formula

$$\rho(x) = Ne^{-2\phi(x)/g^2}$$

whenever ϕ is such that we have $e^{-2\phi(x)/g^2} \in L^1(\mathbb{R}^n)$.

In fact, (16) belongs to a general class of systems that take the form

$$\dot{x} = -\nabla\phi(x) + G(x)\dot{w}$$

where the gradient ∇ is computed with respect to the Riemannian metric given by $(GG^T)^{-1}$. A detailed study of such systems, extended also to degenerate diffusions, is carried out in [6]. An arbitrary compatible system will possess, in addition to gradient terms, certain "skew-symmetric" terms which do not change the steady-state probability distribution (more precisely, these come from vector fields of divergence zero that are everywhere tangential to the equiprobable surfaces). This statement is made precise in [6]; see also Section VII. In fact, all compatible systems naturally fall into the framework of [6] for the case of $\mathbb{R}^n$ with a constant metric. A special class of such systems in $\mathbb{R}^2$ has been described by Rueda in [23].

Example 2: Newton's second law for a nonlinear spring in a viscous fluid in the presence of random external forces may be expressed, with some abuse of notation, by the second-order equation

$$\ddot{x} + a\dot{x} + f(x) = \dot{w} \tag{17}$$

where $\dot{w} = (dw/dt)$ is white noise and $a > 0$. The total energy of the system is $\frac{1}{2}\dot{x}^2 + F(x)$, and a steady-state probability density is

$$\rho(x, \dot{x}) = N[\exp -2a(\tfrac{1}{2}\dot{x}^2 + F(x))]. \tag{18}$$

This formula reflects the facts that the levels of equal energy are also the levels of equal probability in steady state and that the fluctuation introduced by the presence of white noise and the energy dissipation due to the damping term $a\dot{x}$ eventually balance each other. This example is also well known and can be generalized to higher dimensions [12], [16]; see [5] for an application of these ideas to function minimization using simulated annealing.

To understand how the above example fits into our framework, consider the following auxiliary system:

$$\frac{d}{dt}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -\epsilon & 1 \\ 0 & -a \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}\dot{w} + \begin{pmatrix} 0 \\ -1 \end{pmatrix} f(x), \qquad \epsilon > 0. \tag{19}$$

One can check that the compatibility condition (5) holds for (19) with the proportionality constant $\lambda = 2(a+\epsilon) \to 2a$ as $\epsilon \to 0$, which reveals the meaning of the constant $2a$ in the formula (18). If we compute the steady-state density for (19) using the formula (13), and then take the limit as $\epsilon \to 0$, we arrive precisely at (18).

Example 3: As another physical example, consider the circuit shown in Fig. 1. Suppose that its elements are a linear inductor with inductance L, a nonlinear capacitor, and a noisy resistor in the Nyquist-Johnson form (see, e.g., [11]). The Kirchhoff's voltage law reads

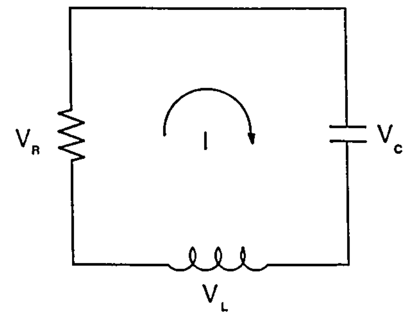

Fig. 1. Electrical circuit of Example 3.

$$V_R + V_C + V_L = 0.$$

For the inductor we have

$$V_L = L\dot{I}.$$

The Nyquist-Johnson resistor model gives

$$V_R = RI + \sqrt{R}\,\dot{w}$$

(this reflects the fluctuation–dissipation equality for an approprate value of the temperature, namely, $1/2$). Finally, for the nonlinear capacitor we have a voltage–charge relationship of the form

$$V_C = f(q_c).$$

Therefore, letting $x_1 = q_c$ (charge) and $x_2 = I$ (current), we obtain the system

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\frac{R}{L}\,x_2 - \frac{\sqrt{R}}{L}\,\dot{w} - \frac{1}{L}\,f(x_1). \end{aligned} \tag{20}$$

The equation for a steady-state density is

$$\frac{R}{2L^2}\,\rho_{x_2x_2} + \frac{1}{L}\,(Rx_2 + f(x_1))\rho_{x_2} - x_2\rho_{x_1} + \frac{R}{L}\,\rho = 0.$$

The energy of the inductor is $\int_0^t IV_L\,dt = \int_0^t LIdI = (L/2)\ x_2^2$. The energy of the capacitor is $\int_0^t IV_C\ dt = \int_0^t f(x_1)\,dx_1 = F(x_1)$. One can verify that a steady-state density is given by

$$\rho(x_1, x_2) = N\exp\left[-2\left((L/2)x_2^2 + F(x_1)\right)\right]$$

(cf. Example 2). Notice that f automatically satisfies the inequality $xf(x) \geq 0$ because it expresses the voltage–charge relationship in the capacitor, hence $e^{-2((L/2)x_2^2+F(x_1))} \in L^1(\mathbb{R}^n)$. The system (20) is compatible in the same limiting sense as (17).

Alternatively, we could represent the nonlinear law for the capacitor as a linear one plus a perturbation, which amounts to letting $f(x) = (1/C)\,x + g(x)$ for some function g and a positive constant C. Further, we could switch to canonical coordinates in which the equipartition of energy property holds (cf. next section). Namely, if we scale the variables, $x_1 \mapsto (1/\sqrt{C})\,x_1,\ x_2 \mapsto \sqrt{L}\,x_2$, then the steady-state variance matrix Q becomes $\frac{1}{2}\ I$, and compatibility can be easily verified using Corollary 3. Moreover, $\lambda = 2$ regardless of the numerical characteristics of the circuit elements. The same method would also work for Example 2.

V. Compatibility from the Statistical Thermodynamics Viewpoint

We are now in position to give an interpretation of the compatibility property on physical grounds. It involves systems that describe the behavior of electrical networks with noisy resistors in Nyquist-Johnson form, as in Example 3 above. If all the resistors are of the same temperature T, the system is called *monotemperaturic*. This concept was mathematically defined in [11], where the authors give a canonical representation for such systems in the form

$$\begin{aligned}\dot{x} &= \left(\Omega - \frac{1}{2T}\, GG^T\right)\, x + G\dot{w} + Du\\ \dot{y} &= -\,D^T x - Fu + \sqrt{2TF}\dot{\nu}.\end{aligned} \tag{21}$$

Here $\Omega = -\Omega^T$, $F = F^T$, and $\dot{w}$ and $\dot{\nu}$ are independent white noise processes. The steady-state variance for (21) upon setting $u = y$ becomes $Q = TI$, so we can say that in steady state all the modes possess equal energy. This property is sometimes referred to as the *equipartition of energy* property; see [11] for a more rigorous justification of this terminology.

In the present framework, certain types of circuits with nonlinear capacitors or inductors are described by equations of the form (1). We claim that by closing the feedback loop in (21) we can obtain a compatible system. Indeed, let $u = y + f(y)$ (assuming single-input, single-output case, otherwise do it for each pair (u_i, y_i)). This yields

$$\begin{aligned}\frac{d}{dt}\begin{pmatrix} x\\ y\end{pmatrix} &= \begin{pmatrix}\Omega - \frac{1}{2T}\, GG^T & D\\ -D^T & -F\end{pmatrix}\begin{pmatrix} x\\ y\end{pmatrix}\\ &\quad + \begin{pmatrix} G & 0\\ 0 & \sqrt{2TF}\end{pmatrix}\begin{pmatrix}\dot{w}\\ \dot{\nu}\end{pmatrix} + \begin{pmatrix} D\\ -F\end{pmatrix} f(y).\end{aligned} \tag{22}$$

We have thus obtained a system that takes the form (15) described in Corollary 3. Summarizing, we can say that compatibility can be thought of as a natural property of monotemperaturic systems with nonlinear reactances. Notice that (1) is more general than (22) since the noise matrix of (1) does not necessarily take a block diagonal form.

Example 4: Consider the circuit shown in Fig. 2.

Let us assume for simplicity that all the resistance and capacitance values are equal to 1, and that the temperature of all the resistors is 1/2. Suppose also that the inductor is nonlinear. Then we can write: $V_1 + V_3 + V_4 - V_2 = 0$, $\dot{V}_1 = I_1$, $\dot{V}_2 = -I_3$, $V_1 = I_2 - I_1 + \dot{w}_1$, $V_2 = I_3 - I_2 + \dot{w}_2$, $V_4 = I_2 + \dot{w}_3$, and $I_2 = \phi + f(\phi)$ where ϕ is the flux in the inductor. The open-loop equations are

$$\begin{aligned}\frac{d}{dt}\begin{pmatrix} V_1\\ V_2\end{pmatrix} &= \begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix} V_1\\ V_2\end{pmatrix}\\ &\quad + \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}\dot{w}_1\\ \dot{w}_2\end{pmatrix} + \begin{pmatrix}1\\ -1\end{pmatrix} I_2\\ \dot{\phi} &= (-1 \quad 1)\begin{pmatrix} V_1\\ V_2\end{pmatrix} - I_2 + \dot{w}_3.\end{aligned}$$

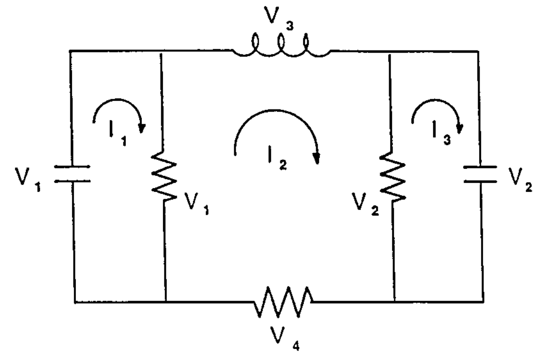

Fig. 2. Electrical circuit of Example 4.

The closed-loop equations are

$$\begin{aligned}\frac{d}{dt}\begin{pmatrix} V_1\\ V_2\\ \phi\end{pmatrix} &= \begin{pmatrix}-1 & 0 & 1\\ 0 & -1 & -1\\ -1 & 1 & -1\end{pmatrix}\begin{pmatrix} V_1\\ V_2\\ \phi\end{pmatrix}\\ &\quad + \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\dot{w}_1\\ \dot{w}_2\\ \dot{w}_3\end{pmatrix} + \begin{pmatrix}1\\ -1\\ -1\end{pmatrix} f(\phi).\end{aligned}$$

VI. Steady-State Densities and Lyapunov Functions

Consider a system excited by a scalar white noise

$$\dot{x} = Ax + g\dot{w} + bf(c^T x).$$

Since by the assumptions made in Section I (A, g) is a controllable pair, in the appropriate basis the linear part of the system takes the standard *controllable companion form,* so that we have

$$\begin{aligned}\frac{d}{dt}\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix} &= \begin{pmatrix} 0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ -p_0 & -p_1 & -p_2 & \cdots & -p_{n-1}\end{pmatrix}\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}\\ &\quad + \begin{pmatrix}0\\ 0\\ \vdots\\ 1\end{pmatrix}\dot{w} + bf(c^T x).\end{aligned}$$

Moreover, if the vectors b and g are proportional, then we can multiply f by a scalar if necessary to arrive at

$$p(D)x + f(c(D)x) = \dot{w}. \tag{23}$$

Here $D := (d/dt)$, and $p(D) = D^n + p_{n-1}D^{n-1} + \cdots + p_1 D + p_0$ and $c(D) = c_{n-1}D^{n-1} + \cdots + c_1 D + c_0$ are polynomials $(D^i x = (d^i x/dt^i))$. The class of systems thus constructed includes (17) and (20) as special cases, and is of considerable interest despite the special form of (23) (see, e.g., [26]).

In this section we will be concerned with formulating conditions on the polynomials p and c under which the system (23) is compatible. We will adopt certain results from [3] regarding the Lur'e problem of absolute stability for the deterministic counterpart of (23)

$$p(D)x + f(c(D)x) = 0. \tag{24}$$

We will assume that $xf(x) > 0$ for all x except $x = 0$, and that either the equation $p(D)x = 0$ is asymptotically stable or $p(D) = Dh(D)$ with $h(D)x = 0$ asymptotically stable. Denoting $Dc(D)$ by $m(D)$, assume also that the function $m(s)/p(s)$ is positive real. Then we can apply the classical *factorization lemma* to conclude that there exists a unique polynomial $r(s)$ with real positive coefficients and no zeros in the right half-plane, such that $\text{Ev}p(s)m(-s) = r(s)r(-s) =: r^{+}(s)r^{-}(s)$ (here "Ev" stands for the even part of a polynomial). We construct a Lyapunov function for (24) as follows:

$$V(x) = \int_0^x \left[p(D)z\, m(D)z - (r^{-}(D)z)^2\right] dt + F(c(D)x) \tag{25}$$

where $(\partial F(c(D)x)/\partial x^{(i)}) = c_i f(c(D)x)$ as before. Obtaining this function is a matter of multiplying both sides of (24) by $m(D)x$, integrating by parts, and completing a square if necessary [3]. In many situations (cf. Example 2) such a Lyapunov function arises naturally as the total energy of the system.

In [3] the problem of absolute stability for (24) is investigated with the aid of the function (25). It can be shown that V is well defined (in particular, the integral in (25) is path-independent), positive definite, and that its derivative along the solutions of (24) is given by

$$\dot{V} = -(r^{-}(D)x)^2. \tag{26}$$

Of course, this expression is in general merely negative semidefinite. To conclude asymptotic stability, LaSalle's principle must be applied, which is essentially what is done in [3].

Now, given a polynomial $p(D)$, let us choose $c(D)$ by setting $c_i = p_{i+1}$ for each even i and $c_i = 0$ for each odd i. Notice that $m(D)$ is then simply the odd part of $p(D)$; therefore $\text{Ev}p(D)m(-D) = m(D)m(-D)$, so we see that $m(s)/p(s)$ is positive real and $r^{-}(D) = m(-D)$. We can also take $m(D)$ to be a constant multiple of $\text{Odd}p(D)$ as in (17), which amounts to a simple modification of the nonlinearity. We will now use the Lyapunov function (25) to arrive at a steady-state density for the stochastic system (23). Assume that $e^{-2aV(x)} \in L^1(\mathbb{R}^n)$.

Theorem 4: Suppose that either $p(D)x = 0$ is asymptotically stable or $p(D) = Dh(D)$ with $h(D)x = 0$ asymptotically stable and that

$$c(D) = \frac{1}{aD}\,\text{Odd}\, p(D) \tag{27}$$

with $a > 0$. Then the function

$$\rho(x) = Ne^{-2aV(x)} \tag{28}$$

is a steady-state probability density associated with the system (23).

Proof: Let us first consider the case when n is even. Using our definition of $c(D)$, it is not difficult to check that the equation for a steady-state probability density can be written as

$$\dot{\rho} = p_{n-1}\rho + \frac{1}{2}\,\frac{\partial^2 \rho}{\partial x_{n-1}^2} \tag{29}$$

where $\dot{\rho}$ stands for the derivative of $\rho(x)$ along the solutions of the deterministic system (24). Plugging into (29) the expression for ρ given by (28) and using (26), we obtain

$$\begin{aligned} 2a(r^{-}(D)x)^2 &= p_{n-1} - ac_{n-2} \\ &\quad + 2a^2(c_{n-2}x_{n-1} + \cdots + c_0 x_1)^2 \\ &= 2(a\, m(D)x)^2 \end{aligned}$$

and the validity of this follows directly from the hypotheses.

The case of odd n is treated similarly; (29) is replaced by

$$\dot{\rho} = (p_{n-1} + \frac{1}{a}\, f'(c(D)x))\rho + \frac{1}{2}\,\frac{\partial^2 \rho}{\partial x_{n-1}^2}$$

and we arrive at the equation

$$a(r^{-}(D)x)^2 = (p(D)x + f(c(D)x) - a\, m(D)x)^2$$

to verify which it remains to use (24) again. □

The formula (27) provides a more concrete interpretation of the compatibility condition (5) applied to systems of the form (23). From the results of Section III it follows that if we want (23) to be compatible, the choice of $c(D)$ is unique up to a constant.

We will now sketch how, in the case when (23) is even-dimensional ($n = 2k$), the above makes contact with the work reported in [27] on steady-state densities for stochastically excited Hamiltonian systems with dissipation which take the form

$$\begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i} \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i} - \sum_{j=1}^{k} a_{ij}\frac{\partial H}{\partial p_j} + \sum_{l=1}^{m} b_{il}\,\dot{w}_l \\ & \quad i = 1, \cdots, k. \end{aligned} \tag{30}$$

In our previous notation, if (27) holds, then c is an even polynomial and the system (23) can be written as

$$\text{Ev}p(D)x + aDc(D)x + f(c(D)x) = \dot{w}. \tag{31}$$

We can think of (31) as the system obtained from

$$\text{Ev}p(D)x + f(c(D)x) = 0 \tag{32}$$

after adding the damping term $aDc(D)x$ and the noise. If p is stable, then the roots of $\text{Ev}p$ are all simple and lie on the imaginary axis, and one has a partial fraction expansion

$$\frac{c(s)}{p(s)} = \sum_{i=1}^{k} \frac{\alpha_i}{s^2 + \beta_i}, \qquad \alpha_i, \beta_i > 0.$$

This implies that there exist coordinates in which (32) takes the Hamiltonian form

$$\begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i} \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i}, \qquad i = 1, \cdots, k \end{aligned}$$

with $H = \frac{1}{2} \sum_{i=1}^{k} (\beta_i q_i^2 + p_i^2) + F(\sum_{i=1}^{k} \sqrt{\alpha_k} q_i)$. It is not hard to see that in these coordinates the original system (30) becomes a special case of (31). In [27], applying techniques from statistical mechanics, Zhu and Yang obtained conditions under which a steady-state density for (30) can be found explicitly. Our presentation here is restricted to the situation when the damping and noise coefficients a_{ij} and b_{il} are constant and the noise is scalar, but the above discussion can be extended to the more general case treated in [27].

The system (23) is a particular case of (1). To establish a connection between the steady-state probability densities and Lyapunov functions in a more general setting, consider the system (1) together with the deterministic one

$$\dot{x} = Ax + bf(c^T x). \tag{33}$$

If the compatibility condition (5) holds, then (33) is equivalent to

$$Q^{-1}A^{-1}\dot{x} = Q^{-1}x + \lambda c f(c^T x). \tag{34}$$

The above analysis and the formula (13) suggest considering the function

$$V(x) = \tfrac{1}{2}\, x^T Q^{-1} x + \lambda F(c^T x)$$

where $(\partial F/\partial x_i) = c_i f(c^T x)$. If $\lambda > 0$ and $xf(x) \geq 0$ for all x, then the function V is positive definite. Its derivative along the solutions of (33) is

$$\begin{aligned}\dot{V} &= \sum_i \frac{\partial V}{\partial x_i}\, \dot{x}_i = \sum_i \left(\sum_j (Q^{-1})_{ij} x_j + \lambda c_i f(c^T x) \right) \dot{x}_i \\ &= \sum_{i,j} \dot{x}_i (Q^{-1}A^{-1})_{ij} \dot{x}_j \\ &= \dot{x}^T (AQ)^{-1} \dot{x} = -\frac{1}{2}\, \dot{x}^T (AQ)^{-1} G G^T ((AQ)^{-1})^T \dot{x}\end{aligned}$$

by virtue of (34) and (4). The last expression is obviously negative semidefinite, which means that V is a Lyapunov function for (33). Thus we see that the steady-state probability densities obtained in Section III are closely related to Lyapunov functions for the corresponding nonrandom systems. As in the case of (24), to conclude asymptotic stability of (33) it is necessary to investigate whether one can apply LaSalle's principle, a question not pursued here.

VII. Existence, Uniqueness, and Convergence

Our goal has been to obtain explicit solutions of (3) under minimal assumptions. In particular, no Lipschitz or other growth requirements have been placed on the function f. Moreover, the diffusion matrix G was not assumed to be nondegenerate—the weaker controllability condition was imposed instead. In this sense, the work reported in this paper serves to *complement* the results available in the literature which allow one to establish existence and uniqueness of steady state and convergence to steady state under additional assumptions. Although for reasons of space we cannot give a complete review of prior work here, we will now provide several examples and a partial reference list.

If the function f is globally Lipschitz, then one can apply the results obtained by Zakai in [26]. Consider the system (1), and denote by L^* the adjoint of the Fokker–Planck operator. Suppose that we have a nonnegative, twice continuously differentiable function $V(x)$ in $\mathbb{R}^n$, which is dominated by a polynomial. Reference [26, Th. 1] can now be formulated as follows: *If there exist numbers $r < \infty$ and $k > 0$ such that $L^* V(x) \leq -k$ for all x satisfying $\|x\| > r$, then the process defined by (1) admits a steady-state probability distribution.* This Lyapunov-like criterion can be used to establish the existence of steady-state probability distributions in the absence of constructive proofs and explicit formulas. This criterion applies to a larger class of systems than that of compatible ones, as will be further illustrated in Section IX.

As for the system (23), in the case of a globally Lipschitz f it satisfies the assumptions of [26, Th. 3]. According to that theorem we therefore conclude that the steady-state density (28) is in fact unique (and so are its special cases considered in Examples 2 and 3). Regarding convergence to this steady-state density, the same theorem of Zakai asserts that for $T \to \infty$ and all $x_0 \in \mathbb{R}^n$ we have

$$\mathcal{P}\left\{ \frac{1}{T} \int_0^T g(x(t))\, dt \to \int_{\mathbb{R}^n} g(x)\rho\, dx \,\middle|\, x(0) = x_0 \right\} = 1$$

where ρ is the steady-state density and g is any real-valued function integrable with respect to the measure $\rho\, dx$.

When the diffusion matrix G is nondegenerate, the process described by (1) can be shown to possess useful properties such as the strong Feller property and, under additional assumptions formulated in [2], recurrence. These properties allow one to apply various results on existence and uniqueness of steady-state probability distributions [1], [14], [17], as well as on convergence to steady state [24], [25]. Another useful concept for establishing uniqueness and convergence is that of a well-behaved solution [12], for which certain growth conditions are required (these conditions will be quite mild, however, in view of the exponential nature of our solutions).

We now outline a constructive method for investigating convergence to steady state, based on the knowledge of the steady-state probability density and on the spectral analysis of the Fokker-Planck operator. If $g_0, g_1, g_2, \cdots$ are the eigenfunctions of the Fokker-Planck operator L corresponding to distinct eigenvalues $\lambda_0, \lambda_1, \lambda_2, \cdots$, then the time-varying solution of the Fokker-Planck equation

$$\frac{\partial \rho}{\partial t} = L\rho \tag{35}$$

with initial condition

$$\rho(0, x) = \sum_{i=0}^{\infty} \alpha_i g_i(x), \qquad \alpha_i \in \mathbb{R}$$

takes the form

$$\rho(t, x) = \sum_{i=0}^{\infty} \alpha_i e^{\lambda_i t} g_i(x).$$

Thus, eigenvalues of L provide information about the convergence of the stochastic process to steady state. For a discussion along these lines, see [4]. In the paper [19], and more recently in [5] and [6], Fokker-Planck operators and their spectral properties were studied with the view toward applications to function minimization. The analysis given below is based on the techniques employed in those references (the subsequent calculations are given in greater detail in [22]).

Suppose that the compatibility condition (5) holds. Consider the function

$$\phi = \tfrac{1}{4}x^TQ^{-1}x + \tfrac{1}{2}\lambda F(c^Tx)$$

(cf. Example 1). Define the vector $\nabla\phi$ by

$$(\nabla\phi)_i = \sum_{j=1}^{n} (GG^T)_{ij}\phi_{x_j}.$$

In view of (4) and (5) we have

$$\nabla\phi = -\tfrac{1}{2}(A + QA^TQ^{-1})x - \tfrac{1}{2}(I + QA^TQ^{-1}A^{-1})bf(c^Tx).$$

The Fokker-Planck operator associated with the system (1) can be written as

$$L = L_{\text{grad}} + L_{\text{skew}}$$

where

$$L_{\text{grad}}\rho = \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\,((\nabla\phi)_i\rho) + \frac{1}{2}\sum_{i,j=1}^{n} (GG^T)_{ij}\rho_{x_ix_j} = \frac{1}{2}\sum_{i,j=1}^{n} (GG^T)_{ij}\,\frac{\partial}{\partial x_i}\left(e^{-2\phi}\,\frac{\partial}{\partial x_j}e^{2\phi}\rho\right)$$

and

$$L_{\text{skew}}\rho = \frac{1}{2}\sum_{i,\,j=1}^{n} \frac{\partial}{\partial x_i}\,[(QA^TQ^{-1} - A)_{ij}x_j\rho] + \frac{1}{2}\sum_{i,\,j=1}^{n} \frac{\partial}{\partial x_i}\,[(QA^TQ^{-1}A^{-1} - I)_{ij}b_jf(c^Tx)\rho].$$

Define a *gauge transform* $\tilde{L}$ of an operator L by $\tilde{L}\rho = e^{2\phi}L(e^{-2\phi}\rho)$. We have

$$\int_{\mathbb{R}^n} \rho\tilde{L}_{\text{grad}}\rho e^{-2\phi}\,dx = -\int_{\mathbb{R}^n} \tfrac{1}{2}\sum_{i,j=1}^{n} (GG^T)_{ij}\rho_{x_i}\rho_{x_j}e^{-2\phi}\,dx \le 0.$$

In the above calculation we used integration by parts. Now consider

$$\int_{\mathbb{R}^n} \rho\tilde{L}_{\text{skew}}\rho e^{-2\phi}\,dx = \int_{\mathbb{R}^n} \tfrac{1}{2}\sum_{i,\,j=1}^{n} \rho[(QA^TQ^{-1} - A)_{ij}x_j]e^{-2\phi}\rho_{x_i}\,dx + \int_{\mathbb{R}^n} \tfrac{1}{2}\sum_{i,\,j=1}^{n} \rho[(QA^TQ^{-1}A^{-1} - I)_{ij}b_jf(c^Tx)] \cdot e^{-2\phi}\rho_{x_i}\,dx + \int_{\mathbb{R}^n} \rho^2L_{\text{skew}}e^{-2\phi}\,dx. \tag{36}$$

Since both L and L_{grad} annihilate the steady-state density $e^{-2\phi}$, it follows that $L_{\text{skew}}e^{-2\phi} = 0$ and so the last integral in the formula (36) is zero. Using integration by parts, we can rewrite the first two integrals in (36) as

$$-\int_{\mathbb{R}^n} \frac{1}{2}\sum_{i,\,j=1}^{n} \rho\frac{\partial}{\partial x_i}\,[(QA^TQ^{-1} - A)_{ij}x_je^{-2\phi}\rho]\,dx - \int_{\mathbb{R}^n} \frac{1}{2}\sum_{i,\,j=1}^{n} \rho\frac{\partial}{\partial x_i}\,[(QA^TQ^{-1}A^{-1} - I)_{ij} \cdot b_jf(c^Tx)e^{-2\phi}\rho]\,dx$$

which implies that

$$\int_{\mathbb{R}^n} \rho\tilde{L}_{\text{skew}}\rho e^{-2\phi}\,dx = -\int_{\mathbb{R}^n} \rho\tilde{L}_{\text{skew}}\rho e^{-2\phi}\,dx = 0.$$

Putting all the above calculations together, we conclude that

$$\int_{\mathbb{R}^n} \rho\tilde{L}\rho e^{-2\phi}\,dx \le 0.$$

Therefore, all eigenvalues of $\tilde{L}$ with eigenfunctions in the space $\{\rho\colon\ e^{-\phi}\rho \in L^2(\mathbb{R}^n)\}$ are nonpositive. This implies that all eigenvalues of L with eigenfunctions in the space $\{\rho\colon\ e^{\phi}\rho \in L^2(\mathbb{R}^n)\}$ are nonpositive. Thus it follows that if the initial probability density $\rho(0, x)$ is a linear combination of such eigenfunctions, then $\rho(t, x)$ approaches steady state. It remains to be seen whether more concrete conclusions can be reached.

Using the above ideas, one can try to obtain specific information about the speed of convergence to steady state, which does not seem possible with the methods described in the references cited earlier. Given the system (1), let us define the *spectral gap* to be

$$\lambda_f = \min_{\rho}\left\{\left(-\int_{\mathbb{R}^n} \rho\tilde{L}\rho e^{-2\phi}\,dx\right)\colon \int_{\mathbb{R}^n} \rho^2e^{-2\phi}\,dx = 1,\ \int_{\mathbb{R}^n} \rho e^{-2\phi}\,dx = 0\right\}.$$

The problem of estimating the spectral gap plays a role in the theory of simulated annealing [19]. In view of the results obtained in [21] which lead to an explicit characterization of the spectral gap for linear systems, it might be useful to relate λ_f to the value of the spectral gap in the case when f is linear. We now show that this can be done, at least in some cases. As an example, consider the familiar spring equation (17) and the corresponding steady-state density (18). The spectral gap is $\lambda_f = \min_\rho\{\frac{1}{2}\int_{\mathbb{R}^2}(\rho_y)^2e^{-y^2-2F(x)}\,dx\,dy\}$ subject to the conditions $\int_{\mathbb{R}^2}\rho^2e^{-y^2-2F(x)}\,dx\,dy = 1$ and $\int_{\mathbb{R}^2}\rho e^{-y^2-2F(x)}\,dx\,dy = 0$. Note that if a function ρ is such that the last condition is satisfied, and if we define $\psi(x, y) := \rho(x, y)e^{\alpha x^2-2F(x)}$ for a certain $\alpha \in \mathbb{R}$, then $\int_{\mathbb{R}^2}\psi e^{-y^2-\alpha x^2}\,dx\,dy = 0$. Now, assume that there exist positive numbers α and ϵ such that $\alpha x^2 \le 2F(x) \le \alpha x^2 + \epsilon$ for all x (geometrically, this means that the graphs of

$f(x)$ and αx must be close enough to each other so that the area between them is finite). Then it is not hard to see that

$$\frac{\int_{\mathbb{R}^2} (\psi_y)^2 e^{-y^2-\alpha x^2}\, dx\, dy}{\left(\int_{\mathbb{R}^2} \psi^2 e^{-y^2-\alpha x^2}\, dx\, dy\right)^2} = \frac{\int_{\mathbb{R}^2} (\rho_y)^2 e^{\alpha x^2 - 2F(x)} e^{-y^2-2F(x)}\, dx\, dy}{\left(\int_{\mathbb{R}^2} \rho^2 e^{\alpha x^2-2F(x)} e^{-y^2-2F(x)}\, dx\, dy\right)^2} \le e^{2\epsilon} \int_{\mathbb{R}^2} (\rho_y)^2 e^{-y^2-2F(x)}\, dx\, dy \le e^{2\epsilon}\lambda_f$$

which implies that $e^{2\epsilon}\lambda_f \ge \lambda_{\alpha x}$, where $\lambda_{\alpha x}$ is the spectral gap for the case $f(x) = \alpha x$. The ease with which we could make the above estimate for this example is due to the fact that in $\tilde{L}_{\text{grad}}$ there is no differentiation with respect to x, the argument of the nonlinearity. Similar estimates hold for any system (1) such that the ith row of G is zero whenever $c_i \neq 0$.

VIII. Compatible Stabilization

Consider the system

$$dx = Ax\, dt + G\, dw - BK\Delta q(x)\, dt \tag{37}$$

where q is a uniform quantizer with sensitivity Δ and an infinite set of values in each direction. In this section we drop the stability assumption a) of Section I and suppose instead that all eigenvalues of $A - BK$ have negative real parts. Rewrite (37) as

$$dx = (A - BK)x\, dt + G\, dw + BKs(x)\, dt \tag{38}$$

where $s(x) := x - \Delta q(x)$. For all x we have

$$\|s(x)\|_\infty \le \Delta/2. \tag{39}$$

Now, assume that the system (38) is compatible in the sense specified at the end of Section II, namely, that $BK = (A - BK)Q_K D$, where $D = \text{diag}(\lambda_1, \cdots, \lambda_n)$ and Q_K is a positive definite symmetric matrix satisfying $(A - BK)Q_K + Q_K(A - BK)^T + GG^T = 0$. Since (38) satisfies the assumption a), we can apply Theorem 2 to obtain a steady-state probability density. Note that the exponential in (13) will automatically belong to $L^1(\mathbb{R}^n)$ by virtue of (39). A possible graph of such a density is sketched in Fig. 3.

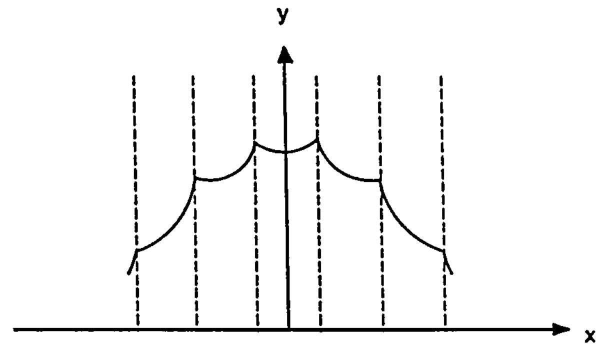

Fig. 3. A steady-state probability density for the system (38).

An interesting question that arises in this context is the following: given the matrices A, G, and B, is it possible to find a matrix K that makes (38) compatible with asymptotically stable nonrandom part? We are not aware of a general answer to this question. What stands in the way of a simultaneous treatment of asymptotic stability and compatibility is the fact that the steady-state variance matrix Q_K associated with the system $dx = (A - BK)x\, dt + G\, dw$ depends on the choice of K and is at the same time needed to verify compatibility of (38). In order to circumvent this difficulty, one might try to develop a procedure which deals with the issues of asymptotic stability and compatibility separately. For instance, even if not all the eigenvalues of A have negative real parts, one might still be able to find a symmetric matrix Q that satisfies the equation (4). In general, this Q need not be positive definite. In view of (12), one must then restrict the search for K to matrices such that BK takes the form AQD with D diagonal. This would then yield a system of the form (38) that possesses a steady-state probability distribution (cf. Remark 1 in Section III).

The above "compatible stabilization" problem is also meaningful for quantized output feedback systems of the form (9). The simplest example (single-input, single-output case) is provided by the familiar nth order equation

$$p(D)x + u = \dot{w} \tag{40}$$

where the control u has to take the form $u = kq(c(D)x)$. We know that a necessary and sufficient condition for this system to be compatible is given by (27). Our choice of the output $c(D)x$ being thus fixed up to a constant, we can rewrite (40) as

$$p(D)x + kc(D)x - ks(c(D)x) = \dot{w}.$$

Note that compatibility is preserved because c is even. Therefore, all we need to do now is choose k that makes the system $p(D)x + kc(D)x = 0$ asymptotically stable. All such values of k can be determined by a straighforward application of the Nyquist criterion.

IX. Non-Compatible Systems and Temperature Bounds

As we have already mentioned, compatibility is not a necessary condition for the existence of a steady-state probability distribution. In the previous sections we have discussed systems for which explicit formulas for steady-state probability densities can be found. It would be interesting to try to develop a perturbation theory that would allow us to obtain specific information about steady-state probability distributions for some systems that are not compatible. We now describe some preliminary results in this direction.

As is well known, if assumption a) of Section I holds, then there exists a positive definite quadratic function $V(x) = x^T Px$ whose derivative along the solutions of $\dot{x} = Ax$ is $-x^T Rx$, with R symmetric positive definite. For the system (1) this implies that outside the ball of radius r

$$\begin{aligned} L^*V(x) &= -x^T Rx + 2x^T Pbf(c^T x) + \text{tr}(PGG^T) \\ &\le -\lambda_{\min}(R)\|x\|^2 \\ &\quad + \max(0, \beta\lambda_{\max}(Pbc^T + cb^T P^T)\|x\|^2) \\ &\quad + \text{tr}(PGG^T) \end{aligned}$$

provided that

$$\alpha z^2 \le z f(z) \le \beta z^2 \quad \text{for all } |z| > r \text{ and some } \beta > \alpha \ge 0 \tag{41}$$

(to arrive at the second term on the right-hand side in the above formula, consider the two cases $x^T Pbf(c^T x) \le 0$ and $x^T Pbf(c^T x) > 0$). Here $\lambda_{\min}$ and $\lambda_{\max}$ stand for the minimal and the maximal eigenvalue, respectively. If we assume that f is globally Lipschitz, then we see that for β small enough there exists a steady-state probability distribution by virtue of [26, Th. 1] cited in Section VII. Observe that the same conclusion holds if $\beta - \alpha$ is small enough, as long as suitable stability conditions are satisfied. Namely, we just have to write $f(c^T x) = (a - \epsilon)c^T x + g(c^T x)$ with a small $\epsilon > 0$ and require that all eigenvalues of $A + (a - \epsilon)bc^T$ have negative real parts.

Example 5: Consider the (noncompatible) second-order system

$$\ddot{x} + f(\dot{x}) + x = \dot{w}. \tag{42}$$

Assuming that (41) holds with $\alpha > 0$, we can recast (42) as $\ddot{x} + \epsilon\dot{x} + g(\dot{x}) + x = \dot{w}$ with $0 < \epsilon < \alpha$. In the above notation, take $P = \left(\begin{smallmatrix} a & 1 \\ 1 & a \end{smallmatrix}\right)$, where $a > 0$. This gives

$$Pbc^T + cb^T P^T = \begin{pmatrix} 0 & -1 \\ -1 & -a \end{pmatrix}$$

with $\lambda_{\max}(Pbc^T + cb^T P^T) = \sqrt{a^2 + 1} - a$, and

$$R = \begin{pmatrix} 2 & \epsilon \\ \epsilon & 2(a\epsilon - 1) \end{pmatrix}.$$

Now, let $\epsilon \to 0$ and $a \to \infty$ in such a way as to have $a\epsilon \to \infty$. Then obviously $\lambda_{\min}(R) \to 2$ and $\lambda_{\max}(Pbc^T + cb^T P^T) \to 0$, which proves the existence of a steady-state probability distribution for all β.

In certain cases it is possible to obtain explicit bounds for the steady-state variance. The corresponding class of systems serves as a natural generalization of monotemperaturic ones. The following observation regarding linear systems provides some motivation for such analysis. Consider the system

$$\dot{x} = (\Omega(t) - G(t)G^T(t))x + F(t)\dot{w} \tag{43}$$

with $\Omega(t)$ skew-symmetric, and assume that there exist two positive numbers T_1 and T_2 (which may be viewed as temperature bounds) such that

$$2T_1 G(t)G(t)^T \le F(t)F(t)^T \le 2T_2 G(t)G(t)^T$$

for all t. The variance equation associated with (43) is

$$\dot{Q} = (\Omega - GG^T)Q + Q(-\Omega - GG^T) + FF^T$$

which implies

$$\begin{aligned}\frac{d}{dt}(T_2 I - Q) = {} & (\Omega - GG^T)(T_2 I - Q) \\ & + (T_2 I - Q)(-\Omega - GG^T) \\ & + 2T_2 GG^T - FF^T\end{aligned}$$

and

$$\begin{aligned}\frac{d}{dt}(Q - T_1 I) = {} & (\Omega - GG^T)(Q - T_1 I) \\ & + (Q - T_1 I)(-\Omega - GG^T) \\ & + FF^T - 2T_1 GG^T.\end{aligned}$$

From this it follows that if $T_1 I \le Q(0) \le T_2 I$, then both $T_2 I - Q$ and $Q - T_1 I$ are nonnegative definite matrices, i.e., $T_1 I \le Q(t) \le T_2 I$ for all t, as one may have expected.

We now single out a class of nonlinear feedback systems perturbed by white noise for which similar bounds can be obtained for certain second moments in steady state. These are single-input, single-output systems of the form

$$\dot{x} = \Omega x - bf(b^T x) + b\dot{w} \tag{44}$$

where $\Omega = -\Omega^T$ (Ω is now time-independent). Let us assume that the condition (41) holds with $\alpha > 0$. In this case it may be interpreted as saying that the temperature of the system (44) is between $1/(2\beta)$ and $1/(2\alpha)$. Rewriting (44) as

$$\dot{x} = (\Omega - bb^T)x - bg(b^T x) + b\dot{w}$$

one easily verifies that it is not compatible unless $\Omega b = 0$. Now, notice that the variance satisfies

$$\frac{d}{dt}\mathcal{E}x^T x = -2\mathcal{E}b^T x f(b^T x) + b^T b$$

(here $\mathcal{E}$ stands for the expectation). Assume for simplicity that $\|b\| = 1$. Provided that the steady-state probability distribution exists, we deduce that in steady state $\mathcal{E}b^T x f(b^T x) = 1/2$ and therefore

$$\frac{1}{2\beta} \le \mathcal{E}(b^T x)^2 \le \frac{1}{2\alpha}$$

by virtue of (41). These are natural bounds imposed by the temperature. Equation (42) may serve as a simple example.

X. Variance Cost and Control Cost

Consider the quantized output feedback system (6), where q is the quantizer with sensitivity $\Delta > 0$ and saturation value 1

$$q(z) = \begin{cases} 1, & \text{if } z \ge \Delta/2 \\ 0, & \text{if } -\Delta/2 \le z < \Delta/2 \\ -1, & \text{if } z < -\Delta/2. \end{cases}$$

Let us assume that the compatibility condition (5) holds with $\lambda > 0$. In the last two sections of the paper we study several steady-state optimal control problems associated with the system (6), with performance criteria formulated in terms of the steady-state probability density. In view of the remarks made in Section VII, solutions to such problems provide a starting point for studying the behavior of the process described by (6) for large times t.

Throughout the rest of the paper, we let $\mathcal{E}$ denote the expectation with respect to the steady-state probability density. The two hyperplanes $\{x : c^T x = \pm\Delta/2\}$ will be referred to as *switching hyperplanes*. We denote by ∇ the gradient computed with respect to the standard Euclidean metric on $\mathbb{R}^n$.

Let us define the *variance cost* by $V := \mathcal{E}(c^T x)^2 = c^T \mathcal{E} xx^T c$. Note that V is invariant under coordinate transformations. We will be interested in the behavior of V considered as a function of the sensitivity Δ. In light of (8) one would expect that for smaller values of Δ the quantized feedback term forces $c^T x$ to become smaller. This idea is formalized in the following statement.

Proposition 5: The variance cost V is a strictly increasing function of the sensitivity Δ.

Proof: The proof is geometrical and uses the explicit formula for the steady-state probability density. If we pick two numbers $\Delta_2 > \Delta_1 > 0$, then in the region $\{x: \Delta_1/2 < c^T x < \Delta_2/2\}$ the steady-state densities corresponding to $\Delta = \Delta_1$ and $\Delta = \Delta_2$ take the form

$$\rho_1(x) = N_1 e^{-(1/2)x^T Q^{-1} x - \lambda c^T x} \tag{45}$$

and

$$\rho_2(x) = N_2 e^{-(1/2)x^T Q^{-1} x} \tag{46}$$

respectively, where N_1 and N_2 are normalization constants. Let us compute the derivatives of ρ_1 and ρ_2 in the direction of the vector $A^{-1}b$. Notice that this vector is transversal to the switching hyperplanes since (5) implies that $c^T A^{-1} b = \lambda c^T Q c > 0$. For ρ_1 we have

$$\begin{aligned}\langle \nabla \rho_1(x), A^{-1}b \rangle &= \langle -(Q^{-1}x + \lambda c), A^{-1}b \rangle \rho_1(x) \\ &= -(x^T Q^{-1} A^{-1} b + \lambda c^T A^{-1} b)\rho_1(x) \\ &= -(\lambda c^T x + \lambda^2 c^T Q c)\rho_1(x).\end{aligned} \tag{47}$$

For ρ_2 we have

$$\begin{aligned}\langle \nabla \rho_2(x), A^{-1}b \rangle &= \langle -Q^{-1}x, A^{-1}b \rangle \rho_2(x) \\ &= -x^T Q^{-1} A^{-1} b \rho_2(x) \\ &= -\lambda c^T x \rho_2(x).\end{aligned} \tag{48}$$

Similar formulas can be obtained for $x \in \{x: -\Delta_2/2 < c^T x < -\Delta_1/2\}$. Now, $\rho_1(x)$ and $\rho_2(x)$ are proportional in the stripe $\{x: |c^T x| < \Delta_1/2\}$. If we assume that $\rho_1(x) \le \rho_2(x)$ there, we arrive at a contradiction. Indeed, (47) and (48) and the fact that $\rho_1(x)$ and $\rho_2(x)$ are also proportional in the region $\{x: |c^T x| > \Delta_2/2\}$ would then imply that $\rho_1(x) < \rho_2(x)$ for $|c^T x| > \Delta_1/2$, and so the requirement

$$\int_{\mathbb{R}^n} \rho_1(x)\, dx = \int_{\mathbb{R}^n} \rho_2(x)\, dx = 1 \tag{49}$$

would be violated. The only remaining possibility is $\rho_1(x) > \rho_2(x)$ for $|c^T x| < \Delta_1/2$. From the above analysis it follows that $\exists \overline{\Delta} \in (\Delta_1, \Delta_2)$ such that $\rho_1(x) = \rho_2(x)$ if $c^T x = \pm\overline{\Delta}/2$, $\rho_1(x) > \rho_2(x)$ if $|c^T x| < \overline{\Delta}/2$, and $\rho_1(x) < \rho_2(x)$ if $|c^T x| > \overline{\Delta}/2$. In fact, (45) and (46) imply that $\overline{\Delta} = 2(\ln N_1 - \ln N_2)/\lambda$. Now it is not difficult to conclude that $V(\Delta_1) < V(\Delta_2)$

$$\begin{aligned}V(\Delta_1) - V(\Delta_2) &= \int_{\mathbb{R}^n} (c^T x)^2 (\rho_1(x) - \rho_2(x))\, dx \\ &< \frac{(\overline{\Delta})^2}{4} \cdot \int_{|c^T x| < \overline{\Delta}/2} (\rho_1(x) - \rho_2(x))\, dx \\ &\quad + \frac{(\overline{\Delta})^2}{4} \cdot \int_{|c^T x| > \overline{\Delta}/2} (\rho_1(x) - \rho_2(x))\, dx \\ &= \frac{(\overline{\Delta})^2}{4} \cdot \int_{\mathbb{R}^n} (\rho_1(x) - \rho_2(x))\, dx = 0\end{aligned}$$

by (49). □

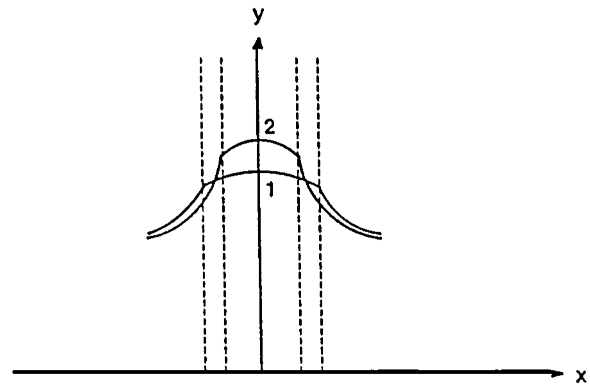

Fig. 4. Relative position of $\rho_1(x)$ and $\rho_2(x)$.

Proposition 5 shows that the optimal control problem that consists of choosing the value of Δ that minimizes V is trivial: the optimal performance is achieved when $\Delta = 0$. Fig. 4 illustrates the previous proof (the cross-section with horizontal coordinate $c^T x$ is shown).

One could also consider the *total variance cost* $\tilde{V} := \mathcal{E} x^T x$ which is invariant under orthogonal coordinate transformations. It turns out that $\tilde{V}$ has the same property of being an increasing function of the sensitivity Δ. This fact can be verified with little work, by choosing one of the coordinate vectors to be parallel to c, using Proposition 5, and integrating by parts in the remaining variables. In fact, the variance $\mathcal{E}(d^T x)^2$ is not affected by the feedback if $d \perp c$.

Another performance criterion that we will consider in this section is the *control cost* $U := \mathcal{P}(|c^T x| > \Delta/2) = \mathcal{E}(q(c^T x))^2$. It may be viewed as the proportion of time during which we actively control the system. This results in certain expenses which we would like to keep low. Clearly, the optimal quantized feedback control strategy consists of moving the switching hyperplanes off to infinity: $\Delta = \infty$ gives $U = 0$. In fact, the following is true.

Proposition 6: The control cost U is a strictly decreasing function of the sensitivity Δ. Δ

The proof is similar to that of Proposition 5; see [22] for details. A nontrivial stochastic optimal control problem consists of determining, for example, the value of Δ that minimizes a quadratic cost functional of the form $L := \alpha V(\Delta) + \sqrt{1-\alpha^2} U(\Delta)$, where $0 < \alpha < 1$. Although this problem does not seem to lend itself to an analytic solution, some numerical results can be obtained.

Example 6: Consider the equation $dx = -x\, dt + dw - q(x)\, dt$, where $x \in \mathbb{R}$ and q is a quantizer with saturation value 1 as before. The diagrams in Figs. 5 and 6 display, respectively, the optimal sensitivity Δ_{opt} as a function of α, and the optimal cost L_{opt} as a function of α (compared with $L = \alpha/2$ for the uncontrolled system $dx = -x\, dt + dw$). Naturally, the quantized feedback control strategy proves most effective when the weight

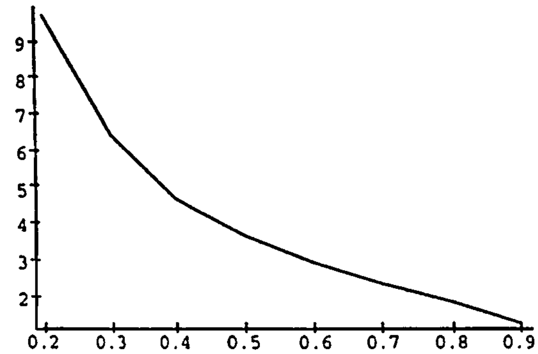

Fig. 5. Optimal sensitivity.

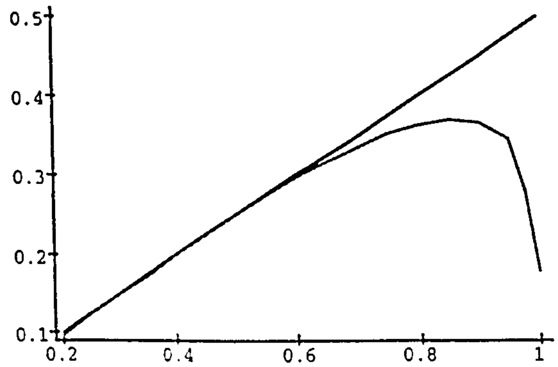

Fig. 6. Optimal cost.

attached to the control cost is small compared to the weight attached to the variance cost: if $\alpha = 1$, choosing $\Delta = 0$ gives $L \approx 0.18$ as opposed to 0.5 for the uncontrolled system.

Since explicit formulas for the steady-state probability densities are available, a number of other optimization problems can be investigated similarly. For example, one could fix Δ and take the proportionality constant λ that appears in the compatibility condition (5) to be the varying parameter (this is meaningful if we can adjust the norm of the vector b). Then it is not hard to show that both the variance cost V and the control cost U decrease as $||b||$ becomes large.

In all of the optimal control problems considered above, the vector c has been fixed. However, the following problem appears interesting: given the system (6) and a fixed sensitivity Δ, choose c so as to minimize the variance cost V, subject to the constraints $||c|| = 1$ and $\lambda = 1$ (b will then be fixed by (5)). Using the formula (8) and an argument similar to the one given in the Proof of Proposition 5, one can show that the optimal c is the eigenvector of Q with the smallest eigenvalue.

XI. Attention Cost

In this section we look at another criterion for evaluating the performance of the system (6). Every time the solution trajectory crosses one of the switching hyperplanes $c^T x = \pm\Delta/2$, we need to communicate to the controller a request to change the control value. This reflects the amount of "attention" needed for implementing the given control law (a similar idea is exploited in [7] in the context of deterministic systems with smooth control functions). One might thus be interested in minimizing the number of such crossings per unit time. Making use of the fact that the steady-state probability density is an even function of x, let us define the *attention cost* to be $C := 2\mathcal{E}C_{\Delta/2}(c^T x)$, where $\mathcal{E}C_u(\xi)$ stands for the mean number of crossings of a level u per unit time by a scalar stochastic process $\xi(t)$.

Since the expectation is computed with respect to the steady-state probability distribution, we may treat the process $c^T x(t)$ as stationary, assuming that it "has reached steady state." Therefore, we may use the celebrated Rice's formula for the mean number of crossings [13]

$$\mathcal{E}C_u(\xi) = \frac{1}{\pi}\sqrt{\frac{-\left.\dfrac{d^2 r(\tau)}{d\tau^2}\right|_{\tau=0}}{r(0)}}\cdot e^{-u^2/2r(0)}$$

where $r(\tau) = \mathcal{E}\xi(t)\xi(t+\tau)$ is the autocorrelation function associated with a stationary stochastic process $\xi(t)$. In our case $r(\tau) = \lim_{t\to\infty} c^T \mathcal{E}x(t)x^T(t+\tau)c$.

Let us first study the following question: when is $\mathcal{E}C_{\Delta/2}(c^T x)$ finite? We will need the following easy statement.

Lemma 7: Assume that $\lim_{\tau\to 0^+} d^2 r(\tau)/d\tau^2$ exists and is finite. Then $\mathcal{E}C_{\Delta/2}(c^T x) < \infty$ if and only if $\lim_{\tau\to 0^+} dr(\tau)/d\tau$ exists and equals zero. □

Proof: Since r is an even function, it is easy to see that in order for its second derivative to exist at the origin it is necessary and sufficient for the first derivative to approach zero as τ approaches zero from the right. □

Example 7: Consider the equation

$$dx = -x\,dt + dw, \qquad x \in \mathbb{R}.$$

We have $\lim_{\tau\to 0^+}(dr(\tau)/d\tau) = -\mathcal{E}x(t)x(t+\tau)|_{\tau=0} = -\mathcal{E}x^2(t) < 0$, so the condition of Lemma 7 is not satisfied.

Let us see whether a control term might help. For the equation

$$dx = -x\,dt + dw - bq(x)\,dt, \qquad b > 0 \tag{50}$$

we have $\lim_{\tau\to 0^+}(dr(\tau)/d\tau) = \mathcal{E}x(t)\,(-x(t) - bq(x(t))) < 0$. Thus we see that it is in fact a nontrivial task to construct a control system with a finite attention cost. The following sufficient condition is a direct consequence of the above developments.

Proposition 8: If the compatibility condition (5) holds and if we have $\lim_{\tau\to 0^+}(dr(\tau)/d\tau) = 0$, then the attention cost associated with the system (6) is well defined and finite.

Consider a general linear stochastic system

$$\begin{aligned} dx &= Ax\,dt + G\,dw \\ y &= c^T x \end{aligned}$$

where $x \in \mathbb{R}^n$ and $w \in \mathbb{R}^m$. For $\tau \geq 0$ we have $\lim_{t\to\infty} \mathcal{E}y(t)y(t+\tau) = \lim_{t\to\infty} c^T\mathcal{E}x(t)x(t+\tau)c = c^T Q e^{A^T\tau}c$, where Q is the steady-state variance matrix satisfying the equation (4). Therefore

$$\lim_{\tau\to 0^+}\frac{dr(\tau)}{d\tau} = c^T Q A^T c = \langle c, AQc\rangle. \tag{51}$$

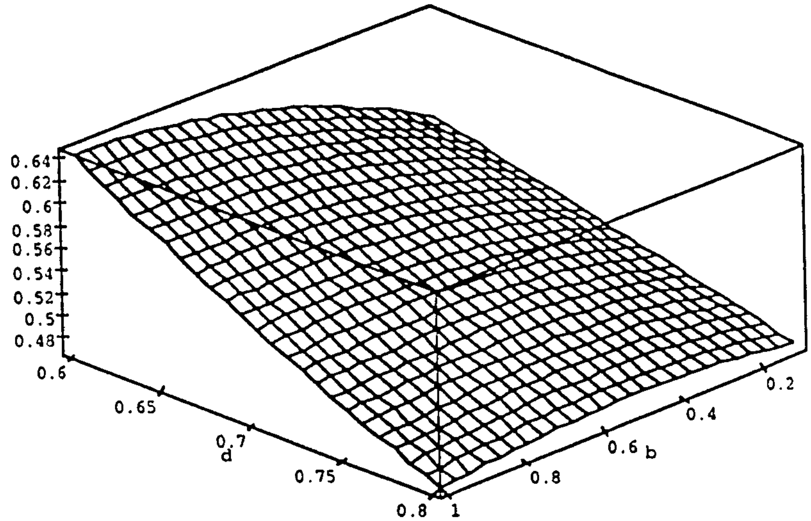

Fig. 7. Attention cost.

Premultiplying (4) by c^T and postmultiplying by c, we obtain

$$2\langle c, AQc\rangle = -\langle G^T c, G^T c\rangle. \tag{52}$$

Thus the condition of Lemma 7 is satisfied if and only if $G^T c = 0$. This means that the attention cost will be finite if all the directions in which the noise can propagate are parallel to the switching hyperplanes, i.e., if the noise "does not contribute directly to the switching hyperplane crossings." On the other hand, if G is a nonsingular $n \times n$ matrix, the attention cost will always be infinite.

The above discussion suggests modifying the equations of Example 7 in the following way: replace $q(x)$ by $q(y)$, where the noise does not enter the equation for y directly. Namely, consider the following pair of equations:

$$\begin{aligned} dx &= -x\ dt + dw \\ dy &= \beta x\ dt - \beta y\ dt \end{aligned} \tag{53}$$

where β is a positive constant (so that the nonrandom part of the system is asymptotically stable). Since the vector $c = \binom{1}{0}$ is orthogonal to the noise vector $g = \binom{1}{0}$, we have $\mathcal{E}C_u(y) < \infty$ by virtue of (51), (52), and Lemma 7.

Now, let us replace (50) by the system

$$\begin{aligned} dx &= -x\ dt + dw - bq(y)\ dt \\ dy &= \beta x\ dt - \beta y\ dt. \end{aligned} \tag{54}$$

It is interesting to observe that (54) is automatically compatible: $g \perp c$ implies $c \perp AQc$ by (52). Therefore, $g \parallel AQc$ since the state space of (54) is $\mathbb{R}^2$. Thus the compatibility condition (5) is satisfied whenever $g \parallel b$, which is indeed the case here. Notice that we did not have to compute the matrix Q to establish compatibility.

We have shown that the mean number of crossings per unit time associated with (53) is finite. Since $q(c^T x)$ is piecewise constant, one might expect that the mean number of crossings for (54) would be finite too, although its value may be different. This is indeed true as we now show. The linear theory applied to (53) yields the steady-state probability density

$$\begin{aligned} \rho(x,y) &= N \exp\left[-(\beta+1)x^2 + 2(\beta+1)xy - \frac{(\beta+1)^2}{\beta}y^2\right] \\ &= N \exp\left[-(\beta+1)(x-y)^2 - \frac{\beta+1}{\beta}y^2\right]. \end{aligned}$$

We have

$$\begin{aligned} \lim_{\tau\to 0^+} &\frac{d}{d\tau}\ \mathcal{E}y(t)y(t+\tau) \\ &= \beta\mathcal{E}(x(t)-y(t))y(t) \\ &= \beta N \iint_{\mathbb{R}^2} (x-y)y \\ &\quad \cdot \exp\left[-(\beta+1)(x-y)^2 - \frac{\beta+1}{\beta}y^2\right]\ dx\ dy \\ &= \beta N \iint_{\mathbb{R}^2} zy \\ &\quad \cdot \exp\left[-(\beta+1)z^2 - \frac{\beta+1}{\beta}y^2\right]\ dz\ dy \\ &= \beta N \int_{-\infty}^{\infty} ze^{-(\beta+1)z^2}\ dz \\ &\quad \cdot \int_{-\infty}^{\infty} ye^{-((\beta+1)/\beta)y^2}\ dy = 0 \end{aligned}$$

where we have used the substitution $z = x - y$. As we have seen in Section II, in the presence of quantized output feedback $q(y)$ we will have an additional term in the exponent, but that term will only depend on y and not on z. Thus the first integral in the last formula will still be zero, and therefore the attention cost for (54) will still be finite.

It is not very difficult to compute this attention cost directly using our knowledge of the steady-state density $\rho(x, y)$ associated with (54). We have

$$\mathcal{E}y^2 = \frac{\beta}{2(\beta+1)} + b^2\mathcal{P}(|y| > \Delta/2) - \frac{\mu\beta b}{\beta+1} \cdot \exp\left[-\frac{(\beta+1)\Delta^2}{4\beta}\right]\sqrt{\frac{\pi}{\beta+1}}$$

where $\mu = \rho(0, 0)$. The attention cost is given by

$$C = \frac{\sqrt{\beta}}{\pi}\sqrt{\frac{\frac{\beta}{2(\beta+1)}}{\mathcal{E}y^2}} \cdot 2e^{-(\Delta^2/8\mathcal{E}y^2)}.$$

Fig. 7 shows the graph of the attention cost C as a function of Δ and b (with β taken to be 1). The corresponding axes are labeled d and b, respectively.

We conclude the paper with an informal discussion of the properties of C revealed by this picture. Suppose that we fix Δ and look at the behavior of C as b varies. We see that C is small for small values of b, because the solution trajectories stay outside the inner stripe $|c^T x| < \Delta/2$ for long periods of time. As b increases, the trajectories return to the inner stripe faster, which results in increasing attention cost. However, as we increase the gain further, the attention cost becomes smaller again and keeps decreasing for bigger values of b. Interestingly, a justification of this latter phenomenon is less obvious. Solving the deterministic counterpart of (53)

$$\dot{x} = -x$$
$$\dot{y} = \beta x - \beta y$$

we have

$$y(t) = \begin{cases} \frac{\beta}{\beta-1}(e^{-t} - e^{-\beta t})x(0) + e^{-\beta t}y(0), & \text{if } \beta \neq 1 \\ te^{-t}x(0) + e^{-t}y(0), & \text{if } \beta = 1. \end{cases} \tag{55}$$

Each of the functions given by the first terms on the right-hand side of both formulas in (55) has a global maximum for a certain $t > 0$, the maximal value being proportional to $x(0)$. Therefore, a solution trajectory which enters the inner stripe far from the y-axis, especially in the I and III quadrants, is likely to leave it again at a later time. However, if we increase b, solution trajectories in those quadrants tend to enter the inner stripe at points closer to the y-axis, so that they are less likely to leave again, hence the decrease in the cost. The smaller the value of Δ is, the bigger value of b is required to achieve the maximal value of C, and the bigger this maximal value is. The reason for this is also seen from the above analysis. On the other hand, if we fix b and increase Δ, the attention cost becomes smaller, by virtue of the fact that the trajectories spend more time in the inner stripe.

One can employ numerical techniques similar to the ones used above to treat stochastic optimal control problems that arise when the cost functional is taken to be a linear combination of the variance cost V, the control cost U, and the attention cost C.

ACKNOWLEDGMENT

The authors would like to thank P. S. Krishnaprasad, D. Ocone, and L. Socha for pointing out useful references and the anonymous reviewers for many helpful comments.

REFERENCES

[1] S. Albeverio, V. Bogachev, and M. Röckner, "On uniqueness of invariant measures for finite- and infinite-dimensional diffusions," *Comm. Pure Appl. Math.*, vol. 52, pp. 325–362, 1999.
[2] R. N. Bhattacharya, "Criteria for recurrence and existence of invariant measures for multidimensional diffusions," *Ann. Prob.*, vol. 6, pp. 541–553, 1978.
[3] R. W. Brockett, "On the stability of nonlinear feedback systems," *IEEE Trans. Appl. Ind.*, vol. 83, pp. 443–449, 1964.
[4] ——, "Lie algebras and Lie groups in control theory," in *Geometric Methods in System Theory*, D. Q. Mayne and R. W. Brockett, Eds. Dordrecht, The Netherlands: Reidel , 1973, pp. 43–82.
[5] ——, "Oscillatory descent for function minimization," in *Current and Future Directions in Applied Mathematics*, M. Alber *et al.*, Eds. Boston, MA: Birkhäuser, 1997, pp. 65–82.
[6] ——, "Notes on stochastic processes on manifolds," in *Systems and Control in the Twenty-First Century*, I. Byrnes *et al.*, Eds. Boston, MA: Birkhäuser, 1997, pp. 75–100.
[7] ——, "Minimum attention control," in *Proc. 36th Conf. Decision and Control*, San Diego, CA, 1997, pp. 2628–2632.
[8] R. W. Brockett and D. Liberzon, "On explicit steady-state solutions of Fokker-Planck equations for a class of nonlinear feedback systems," in *Proc. American Control Conf.*, Philadelphia, PA, 1998, pp. 264–268.
[9] ——, "Quantized feedback systems perturbed by white noise," in *Proc. 37th Conf. Decision and Control*, Tampa, FL, 1998, pp. 1327–1328.
[10] ——, "Quantized feedback stabilization of linear systems," IEEE Trans. Automat. Contr, to be published.
[11] R. W. Brockett and J. C. Willems, "Stochastic control and the second law of thermodynamics," in *Proc. 1978 IEEE Conf. Decision and Control*. New York, 1978, pp. 1007–1011.
[12] T. K. Caughey and F. Ma, "The steady-state response of a class of dynamical systems to stochastic excitation," *ASME J. Appl. Mech.*, vol. 49, pp. 629–632, 1982.
[13] H. Cramér and M. R. Leadbetter, *Stationary and Related Stochastic Processes*. New York: Wiley, 1967.
[14] G. Da Prato and J. Zabczyk, *Ergodicity for Infinite Dimensional Systems*. Cambridge, U.K.: Cambridge Univ. Press, 1996.
[15] D. F. Delchamps, "Stabilizing a linear system with quantized state feedback," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 916–924, 1990.
[16] A. T. Fuller, "Analysis of nonlinear stochastic systems by means of the Fokker-Planck equation," *Int. J. Contr.*, vol. 9, pp. 603–655, 1969.
[17] R. Z. Has'minskii, "Ergodic properties of recurrent diffusion processes and stabilization of the solution to the Cauchy problem for parabolic equations," *Theory Prob. Appl.*, vol. 5, pp. 179–196, 1960.
[18] ——, *Stochastic Stability of Differential Equations*. Alphen aan den Rijn, The Netherlands: Sijthoff & Noordhoff, 1980.
[19] R. A. Holley, S. Kusuoka, and D. W. Stroock, "Asymptotics of the spectral gap with applications to the theory of simulated annealing," *J. Funct. Anal.*, vol. 83, pp. 333–347, 1989.
[20] A. S. Khadr and K. A. Loparo, "Variable structure stochastic systems, conservation laws, and the method of characteristics," in *Proc. Berkeley-Ames Conf. Nonlinear Problems in Control and Fluid Dynamics*, 1983, pp. 315–333.
[21] D. Liberzon and R. W. Brockett, "Spectral analysis of Fokker-Planck and related operators arising from linear stochastic differential equations," SIAM J. Contr, to be published.
[22] D. Liberzon, "Asymptotic properties of nonlinear feedback control systems," Ph.D. dissertation, Dept. Math., Brandeis Univ., Waltham, MA, Feb. 1998.
[23] A. D. Rueda, "A class of two dimensional vector fields which are Fokker-Planck integrable," *Ann. Univ. Sci. Budapest Sect. Math.*, vol. 36, pp. 133–137, 1993.
[24] J. Seidler, "Ergodic behavior of stochastic parabolic equations," *Czechoslovak Math. J.*, vol. 47, pp. 277–316, 1997.
[25] L. Stettner, "Remarks on ergodic conditions for Markov processes on Polish spaces," *Bull. Polish Acad. Sci. Math.*, vol. 42, pp. 103–114, 1994.
[26] M. Zakai, "A Liapunov criterion for the existence of stationary probability distributions for systems perturbed by noise," *SIAM J. Contr.*, vol. 7, pp. 390–397, 1969.

[27] W. Z. Zhu and Y. Q. Yang, "Exact stationary solutions of stochastically excited and dissipated integrable Hamiltonian systems," *ASME J. Appl. Mech.*, vol. 63, pp. 493–500, 1996.

Daniel Liberzon (M'98) was born in Kishinev, former Soviet Union, on April 22, 1973. He received the Ph.D. degree in mathematics from Brandeis University, Waltham, MA, in 1998 (under the supervision of Prof. Roger W. Brockett).

He is currently a postdoctoral associate in the Department of Electrical Engineering at Yale University, New Haven, CT. His research interests include nonlinear control theory, analysis and synthesis of hybrid systems, switching control methods for systems with imprecise measurements and modeling uncertainties, and stochastic differential equations and control.

Dr. Liberzon serves as an Associate Editor on the IEEE Control Systems Society Conference Editorial Board and is a member of SIAM.

Roger W. Brockett (S'62–M'63–SM73–F'74) received the B.S., M.S., and Ph.D. degrees from Case Western Reserve University, Cleveland, OH.

He joined the Department of Electrical Engineering at the Massachusetts Institute of Technology, Cambridge, in 1963 as an Assistant Professor and Ford Foundation Fellow, working in automatic control. In 1969, he was appointed the Gordon McKay Professor of Applied Mathematics in the Division of Applied Sciences at Harvard University, Cambridge, MA. His present position is the An Wang Professor of Electrical Engineering and Computer Sciences at Harvard. His research interests include experimental and theoretical aspects of robotics, including aspects of manipulation, computer control, and sensor data fusion. In addition to being the Associate Director of the Brown-Harvard-MIT Center for Intelligent Control Systems, he also collaborates with colleagues at the University of Maryland through the Maryland-Harvard NSF Engineering Research Center on Systems Engineering.

Dr. Brockett has been involved in the professional activities of the IEEE, SIAM, and AMS, having served on the advisory committees and editorial boards for several groups in these societies. In 1989 he received the American Automatic Control Council's Richard E. Bellman Award, and in 1991, he received the IEEE Control Systems and Engineering Field Award. He is a member of the National Academy of Engineering.