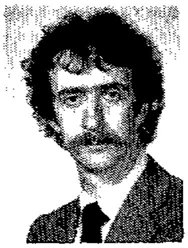
[51] H. Roesler, "Statistical analysis and evaluation of myoelectric signals," *The Control of Upper Extremity Prostheses and Orthoses*, P. Herberts *et al.* Eds. Springfield, IL: Thomas, 1974.
[52] J. K. Salisbury, "Active stiffness control of a manipulator in Cartesian coordinates," presented at the IEEE Conf. on Decision and Contr., NM, 1980.
[53] R. B. Stein and F. Parmiggiani, "Nonlinear summation of contractions in cat muscles: 1. The early depression," *J. General Physiol.*, vol. 78, pp. 277–293, 1981.
[54] E. Taub, I. A. Golberg, and P. Taub, "Deafferentation in monkeys: Pointing at a target without visual feedback," *Exp. Neurol.*, vol. 46, pp. 178–186, 1975.
[55] J. Vrendenbregt and G. Rau, "Surface electromyography in relation to force, muscle length and endurance," *New Developments in EMG and Clinical Neurophysiology*, J. E. Desmedt, Ed. New York: Karger, 1973.
[56] D. R. Wilkie, "The relation between force and velocity in human muscles," *J. Physiol.*, vol. 110, p. 249, 1949.
[57] G. I. Zahalak and S. J. Heyman, "A quantitative evaluation of the frequency response characteristics of active human skeletal muscle in vivo," *J. Biomed. Eng.*, vol. 101, pp. 28–37, 1979.

**Neville Hogan** was born in Dublin, Ireland. He received the Diploma in engineering (with distinction) from Dublin College of Technology, Dublin, Ireland, in 1970 and the S.M. degree in mechanical engineering in 1973, the Mechanical Engineer degree in 1976, and the Ph.D. degree in mechanical engineering in 1977 from the Massachusetts Institute of Technology, Cambridge.

He spent a year in Irish industry as a Product Development and Design Engineer followed by a year as a Research Associate and Lecturer in the Departments of Mechanical Engineering and Psychology at M.I.T. He presently holds an appointment as Associate Professor of Mechanical Engineering at M.I.T., where he teaches graduate and undergraduate subjects in mechanical design, system dynamics, and automatic control, with research activities in the neurophysiology of movement control and the design and control of manipulatory assistive devices and industrial robots.

# Guaranteed Robustness Properties of Multivariable Nonlinear Stochastic Optimal Regulators

JOHN N. TSITSIKLIS AND MICHAEL ATHANS, FELLOW, IEEE

*Abstract*—**We study the robustness of optimal regulators for nonlinear, deterministic and stochastic multiinput dynamical systems, under the assumption that all state variables can be measured. We show that, under mild assumptions, such nonlinear regulators have a guaranteed infinite gain margin; moreover, they have a guaranteed 50 percent gain reduction margin and a 60 degree phase margin in each feedback channel, provided that the system is linear in the control and the penalty to the control is quadratic, thus extending the well-known properties of LQ regulators to nonlinear optimal designs. These results are also valid for infinite horizon, average cost, stochastic optimal control problems.**

## I. Introduction

REGULATOR design for dynamical systems is usually performed on the basis of a nominal model of the plant to be controlled. Modeling errors are unavoidable and, in fact, often desirable because they may result in simpler designs. It is therefore essential that the regulator based on the nominal model is robust; that is, it preserves its qualitative properties (namely, the stability of the closed-loop system) in the face of modeling errors.

The robustness and sensitivity to modeling errors of controlled linear systems has been extensively studied in the past [2], [6]. The robustness (stability margins) of regulators has been traditionally described in terms of gain and phase margins, although more recent approaches [3], [9], [12] focus on the singular values of the return difference or of the inverse return difference matrix.

One of the most appealing features of optimal linear quadratic (LQ) regulators are their guaranteed stability margins. Namely, LQ regulators remain stable when the control gains are multiplied by any number greater than 1/2. They also have guaranteed phase margins of 60 degrees [1], [13], [14], [16]. These results can be obtained directly by appropriately manipulating the associated Riccati equation [13].

A recent paper by Glad [5] has shown that gain margins of optimal regulators for nonlinear systems can be derived from the associated Hamilton–Jacobi–Bellman (HJB) equation, under suitable assumptions. This result ties nicely with the results on LQ regulators because the Riccati equation is a direct consequence of the HJB equation associated with LQ problems. However, the results of [5] are only applicable to single-input, deterministic systems, perturbed by memoryless nonlinearities, thus allowing only derivation of gain margin results; no phase margin results were derived in [5].

In this paper we derive general robustness margins of optimal

Manuscript received March 23, 1983; revised November 21, 1983. Paper recommended by P. R. Kumar, Past Chairman of the Stochastic Control Committee. This work was supported by NASA Ames and Langley Research Centers under Grant NASA/NGL-22-009-124.

J. N. Tsitsiklis was with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139. He is now with the Information Systems Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA 94305.

M. Athans is with the Laboratory for Information Systems, Massachusetts Institute of Technology, Cambridge, MA 02139.

0018-9286/84/0800-0690$01.00 ©1984 IEEE

regulators for multiinput nonlinear systems. Our results are valid for both deterministic and stochastic systems (controlled diffusion processes). In contrast to [5], we allow dynamical (i.e., not just memoryless) perturbations inside the loop and obtain, as a corollary, a generalization of the phase margin results of [13]. In particular, we show (Theorem 3) that the robustness margins of LQ regulators (including the 60 degree phase margin) hold for optimal regulators of any nonlinear plant which is linear in the control, provided that the cost functional contains a quadratic control penalty.

In the stochastic case, we consider two distinct classes of controlled processes. 1) Those for which the state can be steered to an equilibrium point (assumed to be the origin). Such is the case for diffusion processes in which the intensity of the noise decreases to zero as the equilibrium point is approached. We then consider the associated infinite horizon, expected total cost, optimal control problem. 2) Those for which the intensity of the noise is allowed to be everywhere positive. (The LQG problem with perfect observations is an example.) In that case no control law can achieve finite total cost; we consider, however, the associated infinite horizon, expected average cost, optimal control problem. We then derive the same results, provided that stability is now given an appropriate meaning: that no sample path converges to infinity.

We reiterate that the above robustness results are only valid for nonlinear optimal control problems in which all state variables can be measured exactly and can be used in the implementation of the nonlinear feedback regulator. Robustness properties of nonlinear stochastic regulators that arise when only noisy measurements of output variables are available are not addressed in this paper; they remain the subject of future research. Also, we only address robustness issues with respect to plant variations, reflected at the input of the plant. Although more general plant variations are conceivable, this has become a standard way of parameterizing plant uncertainty, at least in the literature on linear systems.

## II. Problem Formulation

*Notation:* Throughout this paper, scalar functions will be indicated by lower case letters; vector functions by lower case bold face letters; matrix functions by upper case letters. For any vector function $\mathbf{f}$ we will use subscripts (e.g., $f_i$) to denote its scalar components. For any scalar function $f$ of a vector input $\mathbf{x}$, we let $\partial f/\partial \mathbf{x}$ denote the transpose of the gradient of $f$ (a row vector).

### *Case A — Deterministic Optimal Control*

Consider the controlled deterministic system

$$\frac{d\mathbf{x}}{dt}(t)=\mathbf{f}(\mathbf{x}(t),\mathbf{u}(t)); \qquad \mathbf{x}(0)=\mathbf{x}_0, \tag{1}$$

where $\mathbf{x},\mathbf{u}$ are $n$- and $m$-dimensional state and control vectors, respectively, and $\mathbf{f}$ is a continuous function from $R^{n+m}$ into $R^n$ such that $\mathbf{f}(0,0)=0$. A control law is a measurable function $\mathbf{k}$: $R^n \mapsto R^m$ such that the closed-loop equation

$$\frac{d\mathbf{x}}{dt}(t)=\mathbf{f}(\mathbf{x}(t),\mathbf{k}(\mathbf{x}(t))); \qquad \mathbf{x}(0)=\mathbf{x}_0, \tag{2}$$

has a unique solution, for all $\mathbf{x}_0 \in R^n$. (If $\mathbf{k}(\cdot)$ is not continuous some more care may be needed in defining what is meant by a solution to (2); see [4].) Let $l$: $R^n \mapsto R$ and $h$: $R^m \mapsto R$ be nonnegative measurable functions denoting the penalties to the state and the control, respectively, satisfying $l(0)=h(0)=0$. We consider the performance criterion

$$J_1=\int_0^\infty \left(l(\mathbf{x}(t))+h(\mathbf{u}(t))\right)dt. \tag{3}$$

The general dynamic programming conditions for optimality for such control problems are well known and easy to establish formally. However, for our purposes, we do not need to concern ourselves with the particular assumptions that can guarantee existence of optimal control laws or that the Hamilton–Jacobi–Bellman (HJB) equation is satisfied. Such issues are treated, for example, in [4] for finite horizon problems. We will assume instead that the data of the control problem are sufficiently well-behaved to guarantee that no complication will arise. (For certain types of control problems, Assumption 1a below may fail to hold, for example if $V$ is not everywhere differentiable; however, it seems that robustness results may be proved even if (4) fails to hold on some subset of the state space, e.g., on the subset where $V$ is not differentiable.) In particular, we assume the following.

*Assumption 1a:* There exists an optimal control law $\mathbf{k}(\cdot)$. Moreover, the optimal cost-to-go (value) function $V$: $R^n \mapsto R$ is continuously differentiable and satisfies the HJB equation

$$\begin{aligned} 0&=\frac{\partial V}{\partial \mathbf{x}}(\mathbf{x})\cdot\mathbf{f}(\mathbf{x},\mathbf{k}(\mathbf{x}))+h(\mathbf{k}(\mathbf{x}))+l(\mathbf{x})\\ &\leq \frac{\partial V}{\partial \mathbf{x}}(\mathbf{x})\cdot\mathbf{f}(\mathbf{x},\mathbf{u})+h(\mathbf{u})+l(\mathbf{x}), \qquad \forall \mathbf{x}\in R^n,\ \forall \mathbf{u}\in R^m. \end{aligned} \tag{4}$$

Finally, $V(\mathbf{x})>0$, $\forall \mathbf{x}\neq 0$ and $\liminf_{\|\mathbf{x}\|\to\infty} V(\mathbf{x})>0$.

### *Case B — Stochastic Optimal Control: Total Cost*

Consider the perfectly observed controlled diffusion process

$$d\mathbf{x}(t)=\mathbf{f}(\mathbf{x}(t),\mathbf{u}(t))\,dt+\Sigma(\mathbf{x}(t))\,d\mathbf{w}(t); \qquad \mathbf{x}(0)=\mathbf{x}_0, \tag{5}$$

where $\mathbf{x},\mathbf{u},\mathbf{f}$ are as in Case A, except that $\mathbf{f}$ is now allowed to be any measurable function; $\Sigma(\mathbf{x})$ is a measurable $n\times n$ matrix function, $\mathbf{w}(t)$ is a standard $n$-dimensional Brownian motion, and $\mathbf{x}_0$ is the initial state. We also assume that $\mathbf{w}(t)$ is defined on some probability space $(\Omega,\mathscr{F},\mathscr{P})$ and we denote by $\mathscr{F}_t$ the smallest $\sigma$-field in $\mathscr{F}$ such that $\mathbf{w}(\tau)$ is $\mathscr{F}_t$-measurable, for all $\tau\leq t$.

A control law is a measurable function $\mathbf{k}$: $R^n \mapsto R^m$ such that the stochastic differential equation

$$d\mathbf{x}(t)=\mathbf{f}(\mathbf{x}(t),\mathbf{k}(\mathbf{x}(t)))\,dt+\Sigma(\mathbf{x}(t))\,d\mathbf{w}(t); \qquad \mathbf{x}(0)=\mathbf{x}_0 \tag{6}$$

has a unique solution in the Ito sense [15]. We consider the performance criterion

$$J_2=\lim_{T\to\infty} E\left[\int_0^T \left(l(\mathbf{x}(t))+h(\mathbf{u}(t))\right)dt\right] \tag{7}$$

where $l(\cdot)$ and $h(\cdot)$ are as for Case A. Let $\Sigma^T(\mathbf{x})$ denote the transpose of $\Sigma(\mathbf{x})$. Let $A(\mathbf{x})=(1/2)\Sigma(\mathbf{x})\Sigma^T(\mathbf{x})$ and $a_{ij}(\mathbf{x})$ be the $i,j$th entry of $A(\mathbf{x})$. We define a differential operator $L^{\mathbf{u}}$ by

$$L^{\mathbf{u}}=\sum_{i=1}^n f_i(\mathbf{x},\mathbf{u})\frac{\partial}{\partial x_i}+\sum_{i,j=1}^n a_{ij}(\mathbf{x})\frac{\partial^2}{\partial x_i\,\partial x_j}. \tag{8}$$

As in the deterministic case, we will assume the following.

*Assumption 1b:* There exists an optimal control law $\mathbf{k}(\cdot)$. Moreover, the optimal cost-to-go (value) function $V$: $R^n \to R$ is twice continuously differentiable and satisfies the HJB equation

$$0 = (L^{k(x)}V)(x) + h(k(x)) + l(x) \leqslant (L^u V)(x) + h(u) + l(x), \qquad \forall x \in R^n,\ \forall u \in R^m. \tag{9}$$

Finally, $V(x) > 0$, $\forall x \neq 0$ and $\liminf_{\|x\| \to \infty} V(x) > 0$.

Conditions under which Assumption 1b is satisfied may be obtained along the lines of [4], [7], [10] and they need not concern us here.

*Case C — Stochastic Optimal Control: Average Cost*

Let everything be as in Case B above, except that the performance criterion is modified to be

$$J_3 = \lim_{T \to \infty} \frac{1}{T} E\left[\int_0^T l(x(t)) + h(u(t))\, dt\right]. \tag{10}$$

We also require that $\{x\colon l(x) < c\}$ is bounded, for any constant $c$. Assumption 1b must then be modified as follows [8].

*Assumption 1c:* There exists an optimal control law $k(\cdot)$, a constant $g$, and a function $V\colon R^n \to R$ which is twice continuously differentiable and satisfies the HJB equation

$$g = (L^{k(x)}V)(x) + h(k(x)) + l(x) \leqslant (L^u V)(x) + h(u) + l(x), \qquad \forall x \in R^n,\ \forall u \in R^m. \tag{11}$$

Suppose that the optimal regulator $u(t) = k(x(t))$, for any of the problems A, B, or C above, is perturbed to $u(t) = \Phi(k(x(\cdot)))(t)$, as in Fig. 1. We are interested in the stability of the new closed-loop system under suitable assumptions on $\Phi$. The perturbation $\Phi$ may be simply a memoryless nonlinearity, in which case we can make statements about the gain margins of the optimal regulator. It can also be a causal dynamical operator (e.g., a linear time-invariant system); in particular if $\Phi$ corresponds to a pure phase shift, we can make statements about the phase margins of the optimal regulator.

We now proceed to define the class of admissible perturbations $\Phi$. Let $k(\cdot)$ be a control law for any of the problems A, B, or C. Let $\mathscr{M}$ denote the set of measurable $m$-dimensional time functions from $[0,\infty)$ into $R$. An *admissible perturbation* $\Phi$ of $k(\cdot)$ is a map from $\mathscr{M}$ into $\mathscr{M}$ such that:

*Case A (Deterministic Systems):* There exists some $u(\cdot) \in \mathscr{M}$ such that:

i) the differential equation (1) has a unique solution $x(\cdot)$,

ii) $u(t) = \Phi(k(x(\cdot)))(t)$, $\forall t$.

*Cases B and C (Stochastic Systems):* There exists a measurable stochastic process $u(t)$ defined on $(\Omega, \mathscr{F}, \mathscr{P})$ such that:

i) $u(t)$ is adapted to $\{\mathscr{F}_t\}$,

ii) the stochastic differential equation (5) has a unique solution $x(\cdot)$,

iii) for any sample path, $u(t) = \Phi(k(x(\cdot)))(t)$.

*Assumption 2:* Let $\mathscr{M}_0 = \{u(\cdot) \in \mathscr{M}\colon \int_0^\infty h(u(\tau))\, d\tau < \infty\}$. Then, $\Phi$ maps $\mathscr{M}_0$ into $\mathscr{M}_0$. (For example, if $h$ is a quadratic function, $\Phi$ must map $L_2$ into $L_2$.)

The solution $x(t)$ of either (1) or (5), when $u(t)$ is given as in the above definition, will be called the "trajectory of the perturbed closed-loop system."

## III. Main Results

Our first result is a multiloop generalization of [5, Theorem 3] which also covers stochastic control problems. It shows that optimal regulators have an infinite gain margin, provided that the following condition is satisfied.

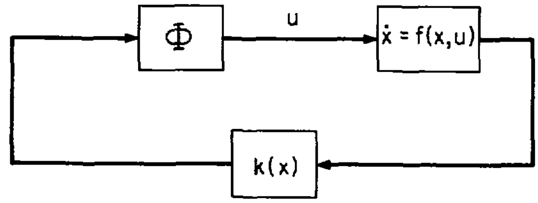



Fig. 1. The perturbed closed-loop system.

*Assumption 3:* i) $f(x,u)$ is differentiable with respect to $u$, for any fixed $x \in R^n$. ii) For any fixed $x \in R^n$, $a \in R^n$, $i \in \{1,\cdots,m\}$, either

$$\frac{\partial}{\partial u_i}\left(a^T \cdot f(x,u)\right) \geqslant 0, \qquad \forall u \in R^m, \tag{12}$$

or

$$\frac{\partial}{\partial u_i}\left(a^T \cdot f(x,u)\right) \leqslant 0, \qquad \forall u \in R^m. \tag{13}$$

*Assumption 4:* For each $t$, there exist functions $\phi_i(\cdot,t)\colon R \to R$ such that

$$\Phi(u(\cdot))(t) = \left(\phi_1(u_1(t),t),\cdots,\phi_m(u_m(t),t)\right)^T, \qquad \forall u(\cdot) \in \mathscr{M}. \tag{14}$$

Moreover, these functions satisfy, for each $t$, the sector condition (Fig. 2)

$$c^2 \leqslant c\phi_i(c,t), \qquad \forall c \in R,\ \forall i. \tag{15}$$

In other words, the perturbation $\Phi$ corresponds to a memoryless nonlinearity and, in particular, to a gain increase.

*Theorem 1:* Consider the optimal control problems A, B, C and let Assumptions 1a, 1b, 1c, respectively, as well as Assumption 3, hold and suppose that $h(u) = \sum_{i=1}^m h_i(u_i)$, for appropriate scalar functions $h_i$ such that $h_i(u_i) > 0$, $\forall u_i \neq 0$. Let $\Phi$ be an admissible perturbation of a corresponding optimal control law, satisfying Assumptions 2 and 4, and let $x(t)$ denote the trajectory of the perturbed closed-loop system. Then,

*Case A (Deterministic Problems):* $\lim_{t\to\infty} x(t) = 0$.

*Case B (Stochastic Total Cost Problems):* $\lim_{t\to\infty} x(t) = 0$, almost surely.

*Case C (Stochastic Average Cost Problems):* No sample path converges to infinity, almost surely.

Thus, in all cases the perturbed nonlinear closed-loop system remains stable.

*Proof:* All proofs can be found in the Appendix.

Let us comment briefly on the meaning of stability for the average cost case. Our result states that no sample path converges to infinity, which is equivalent to saying that for every sample path there exists a bounded set $K$ and a sequence $t_n$ converging to infinity, such that $x(t_n) \in K$, $\forall n$. This does not mean that sample paths are bounded. For typical average cost problems (e.g., for the LQG problem with perfect observations), sample paths are unbounded, almost surely, even if an optimal control law is used.

We now discuss the crucial Assumption 3. Theorem 1 remains true even if $f(x,\cdot)$ is not differentiable, provided that Assumption 3ii) is appropriately modified, as in [5] (although the more general version is more obscure). However, the proof of Theorem 1 reveals that it cannot be significantly further weakened. Assumption 3 essentially guarantees that the (expected) direction of motion is still a descent direction (with respect to the value function $V$) under an arbitrary gain increase. Given the importance of Theorem 1, it is a natural question to find particular cases for which Assumption 3 holds. Glad [5] provides the

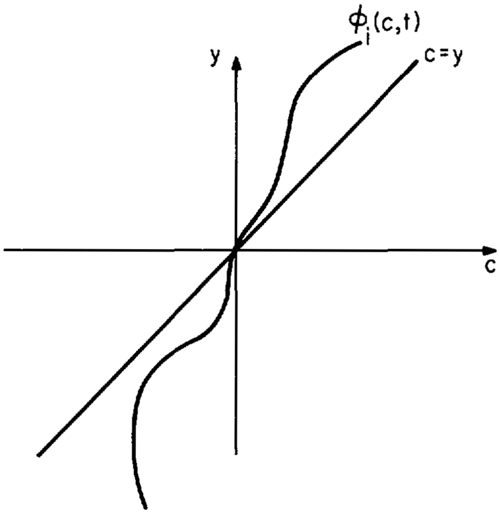



Fig. 2. A memoryless gain increase.

example (for the single input case)

$$f(x,u)=f^1(x)+b(x)f^2(x,u) \tag{16}$$

where $f^1$: $R^n \to R^n$, $b$: $R^n \mapsto R^n$, $f^2$: $R^{n+1} \mapsto R$ and where $f^2(x,\cdot)$ is monotonic in $u$, for any fixed $x$. Interestingly enough, the above example covers *all* cases allowed by Assumption 3 and a similar characterization can be also obtained for the multiinput case. This is the subject of the next theorem, in which we assume that $f$ is twice continuously differentiable with respect to $u$ because this allows a significant simplification of the proof.

*Theorem 2:* Let $f$ be twice continuously differentiable with respect to $u$, for any fixed $x$. Then, $f$ satisfies Assumption 3ii) if and only if, for any fixed $x$, it is of the form

$$f(u)=b^0+\sum_{k=1}^{q} b^k f^k(u) \tag{17}$$

where $f^k$: $R^m \mapsto R$. Moreover, for any component $u_i$ of $u$, at most one of the scalar functions $f^k$ may depend on $u_i$. Finally, each function $f^k$ is either increasing in $u_i$, for all $u$, or decreasing in $u_i$, for all $u$. (That is the scalar functions $f^k$ satisfy themselves Assumption 3. However, the way that components are split to form the sum in (17) may change with $x$.)

As in [5] more assumptions on the dynamics are needed to obtain more specific robustness margins. In what follows we assume that the dynamics are linear in the control.

*Assumption 5:* $f(x,u)=f^0(x)+F(x)u$, where $f^0$: $R^n \mapsto R^n$ and $F(x)$ is a $n\times m$ matrix function, for each $x\in R^n$.

The next assumption describes the set of perturbations $\Phi$ that will be allowed. It may seem counterintuitive as stated below in its full generality. In fact, it is a generalization of the conditions imposed in either [5] or [13] as will be shown later.

*Assumption 6:* i) $h(\cdot)$ is continuously differentiable. ii) There exists some $\epsilon>0$ such that for any measurable $m$-dimensional time function $u(\cdot)$, and for any $t\geqslant 0$,

$$\int_0^t \left[\frac{\partial h}{\partial u}(u(\tau))\cdot[\Phi(u(\cdot))(\tau)-u(\tau)]+(1-\epsilon)h(u(\tau))\right]d\tau \geqslant 0. \tag{18}$$

*Theorem 3:* Consider the optimal control problems A, B, C and let Assumptions 1a, 1b, 1c, respectively, as well as Assumption 5, hold. Let $\Phi$ be an admissible perturbation of a corresponding optimal control law, satisfying Assumptions 2 and 6, and let $x(t)$ denote the trajectory of the perturbed closed-loop system. Then:

*Case A (Deterministic Problems):* $\lim_{t\to\infty}x(t)=0$.

*Case B (Stochastic Total Cost Problems):* $\lim_{t\to\infty}x(t)=0$, almost surely.

*Case C (Stochastic Average Cost Problems):* No sample path converges to infinity, almost surely.

Thus, in all cases the perturbed nonlinear closed-loop system remains stable.

In order to apply Theorem 3, one mainly needs to verify that Assumption 6 holds. This is done below for certain particular problems. Proposition 1 shows that the robustness margins of LQ regulators generalize to nonlinear systems which are linear in the control and in which the penalty to the control is quadratic.

*Proposition 1:* Suppose that $h(u)=\sum_{i=1}^m r_i u_i^2$ ($r_i>0$) and let $\Phi$ be a linear time invariant system with diagonal transfer matrix whose nonzero entries $\phi_i$ are proper, stable rational functions, and, for some $\epsilon>0$, $\mathrm{Re}[\phi_i(j\omega)]\geqslant 1/2+\epsilon$, $\forall\omega$. Then, Assumption 6 holds (possibly with a different $\epsilon$). The condition $\mathrm{Re}[\phi_i(j\omega)]\geqslant 1/2+\epsilon$ is satisfied, in particular, if for some $\epsilon>0$:

i) $\Phi$ is a memoryless gain, larger than $1/2+\epsilon$, in each channel, or

ii) $\Phi$ is a pure phase shift, smaller than $60-\epsilon$ degrees at all frequencies, in each channel.

Proposition 1 showed that Theorem 3 generalizes the LQ gain and phase margin results of [13]. The next proposition shows that the same theorem generalizes the single-input results of [5] as well.

*Proposition 2:* Suppose that $h(u)=\sum_{i=1}^m h_i(u_i)$ and let $\Phi$ be a memoryless nonlinearity such that

$$\Phi(u(\cdot))(t)=(\phi_1(u_1(t),t),\cdots,\phi_m(u_m(t),t))^T \tag{19}$$

and, for some $\epsilon>0$,

$$u_i\phi_i(u_i)\geqslant u_i^2-(1-\epsilon)\frac{u_ih_i(u_i)}{\dfrac{\partial h_i}{\partial u_i}(u_i)}, \qquad \forall u_i\in R,\ \forall i. \tag{20}$$

Moreover, assume that $u_i(\partial h_i/\partial u_i)(u_i)>0$, $\forall u_i\neq 0$. Then Assumption 6 is satisfied.

Proposition 2 may provide us with gain reduction margin results. As an application, let $h_i(u_i)=u_i^{2n}$, for some positive integer $n$. Inequality (20) becomes

$$u_i\phi_i(u_i)\geqslant u_i^2\left(1-\frac{1-\epsilon}{2n}\right) \tag{21}$$

which shows that the stronger we penalize large inputs ($n$ large), the worse become the gain reduction margins, as should be expected.

## IV. Conclusions

This paper demonstrates that under suitable assumptions nonlinear optimal multiinput deterministic or stochastic dynamic systems have certain guaranteed robustness properties, which may be expressed as guaranteed gain and phase margins. These properties generalize the known robustness results of optimal regulators for linear systems with respect to quadratic performance criteria. In particular it is shown that if the nonlinear dynamic system is *linear* in the control variables and there is a *quadratic* penalty on the control variables in the associated cost functional, then the resulting nonlinear feedback design has a guaranteed infinite positive gain margin, a $-6$ dB gain reduction margin, and a $\pm 60$ degree phase margin property.

Such guaranteed robustness properties are obtained from the Hamilton–Jacobi–Bellman equation associated with the nonlinear optimal control problems.

## APPENDIX

This Appendix contains all proofs for Section III.

*Lemma 1:*

*Case A:* Given some $x(0) \in R^n$ and a time function $u(t)$, let $x(t)$ be the corresponding solution of (1), assuming that it exists. Assume that $\int_0^\infty [l(x(\tau)) + h(u(\tau))]\, d\tau < \infty$. Then, $\lim_{t\to\infty} x(t) = 0$.

*Case B:* Given some $x(0)$ and a stochastic process $u(t)$, adapted to $\mathscr{F}_t$, let $x(t)$ be the corresponding solution of (5), assuming that it exists. Assume that $\int_0^\infty [l(x(\tau)) + h(u(\tau))]\, d\tau < \infty$, almost surely, and that $\sup_t V(x(t)) < \infty$, almost surely. Then $\lim_{t\to\infty} x(t) = 0$, almost surely.

*Proof:*

*Case A:* Let $\hat V(t) = \int_t^\infty (l(x(\tau)) + h(u(\tau)))\, d\tau$. Then, $\hat V(t) \geq V(x(t))$, since $V$ is the optimal cost-to-go function. Clearly, $\lim_{t\to\infty} \hat V(t) = 0$, which implies that $\lim_{t\to\infty} V(x(t)) = 0$. By Assumption 1a, it follows that $\lim_{t\to\infty} x(t) = 0$.

*Case B:* a) We first consider the case where

$$E\left[\int_0^\infty (l(x(\tau)) + h(u(\tau)))\, d\tau\right] < \infty$$

also holds. With $\hat V(t)$ defined as for Case A, it is easy to see that $\lim_{t\to\infty} E[\hat V(t)|\mathscr{F}_t] = 0$, a.s. Moreover, the definition of $V$ implies that $E[\hat V(t)|\mathscr{F}_t] \geq V(x(t))$, which shows that $V(x(t))$ converges to zero and, using Assumption 1b, $x(t)$ must converge to the origin.

b) We now consider the general case. Given the initial state $x(0)$ and some $M \geq 0$, $N \geq 0$ let

$$T_{MN} = \inf\left\{ t \geq 0 \colon \int_0^t (l(x(\tau)) + h(u(\tau)))\, d\tau \geq N \text{ or } V(x(t)) \geq M \right\}.$$

If the above set is empty, let $T_{MN} = \infty$. We now define a new control law $\hat u$ by

$$\hat u(t) = \begin{cases} u(t), & t < T_{MN} \\ k(x(t)), & t \geq T_{MN}, \end{cases}$$

where $k(\cdot)$ is an optimal control law and let $\hat x(\cdot)$ be the trajectory that results when $\hat u$ is used. Then,

$$\begin{aligned} &E\left[\int_0^\infty (l(\hat x(\tau)) + h(\hat u(\tau)))\, d\tau\right] \\ &= E\left[\int_0^{T_{MN}} (l(x(\tau)) + h(u(\tau)))\, d\tau\right] + E[V(x(T_{MN}))] \\ &\leq M + N. \end{aligned}$$

Therefore, by part a), $\lim_{t\to\infty} \hat x(t) = 0$, almost surely. Let $\Omega_{MN} = \{\omega \in \Omega \colon T_{MN} = \infty\}$. For all $\omega \in \Omega_{MN}$ we have $\hat x(t) = x(t)$, $\forall t$. Hence, $\lim_{t\to\infty} x(t) = 0$, for almost all $\omega \in \Omega_{MN}$, $\forall M, N$. On the other hand, the assumptions of the lemma imply that almost all $\omega \in \Omega$ also belong to $\Omega_{MN}$, for some $M, N$, and the desired result follows. ∎

*Proof of Theorem 1:* We follow the approach of [5]. Fix some $x \in R^n$. Then, by Assumption 3, $\partial V/\partial x(x) \cdot f(x,u)$ is either increasing or decreasing, as a function of $u_i$. Assume it is increasing. From either (4), (9), or (11) corresponding to Cases A, B, C, respectively, we obtain

$$\frac{\partial V}{\partial x}(x) \cdot f(x, k(x)) + h(k(x)) \leq \frac{\partial V}{\partial x}(x) \cdot f(x,u) + h(u), \tag{A1}$$

for all $u$ that differ from $k(x)$ in the $i$th component only. It follows that $k_i(x) \leq 0$ because otherwise

$$h_i(k_i(x)) > h_i(0)$$

and

$$\frac{\partial V}{\partial x}(x) \cdot f(x, k(x)) \geq \frac{\partial V}{\partial x}(x) \cdot f(x, k^*(x))$$

where $k_i^*(x) = 0$, $k_j^*(x) = k_j(x)$, $j \neq i$, which would contradict (A1). Similarly, we conclude that $k_i(x) \geq 0$, whenever $\partial V/\partial x(x) \cdot f(x,u)$ is decreasing as a function of $u_i$.

Assumption 4 implies that $\phi_i(k_i(x), t) \geq k_i(x)$ whenever $k_i(x) \geq 0$ and $\phi_i(k_i(x), t) \leq k_i(x)$, otherwise. Together with the preceding discussion we conclude that

$$\frac{\partial V}{\partial x}(x) \cdot f(x, \phi(k(x), t)) \leq \frac{\partial V}{\partial x}(x) \cdot f(x, k(x)). \tag{A2}$$

From now on, let $x(t)$ denote the trajectory of the perturbed closed-loop system and let $u(t) = \phi(k(x(t)))$.

*Case A (Deterministic Problems):* From inequality (A2) and (4) we obtain

$$\begin{aligned} \frac{dV}{dt}(x(t)) &= \frac{\partial V}{\partial x}(x(t)) \cdot f(x(t), \phi(k(x), t)) \\ &\leq \frac{\partial V}{\partial x}(x(t)) \cdot f(x(t), k(x(t))) \\ &= -l(x(t)) - h(k(x(t))) \leq 0. \end{aligned} \tag{A3}$$

Integrating (A3), we obtain

$$\int_0^t [l(x(\tau)) + h(k(x(\tau)))]\, d\tau \leq V(x(0)) - V(x(t)) \leq V(x(0))$$

and therefore, $\int_0^\infty [l(x(\tau)) + h(k(x(\tau)))]\, d\tau < \infty$. In view of Assumption 2, the last inequality also gives $\int_0^\infty h(u(\tau))\, d\tau < \infty$. The desired result then follows from Lemma 1.

*Case B (Stochastic Total Cost Problems):* From inequality (A2) and (9) we obtain

$$\begin{aligned} (L^{u(t)}V)(x(t)) &\leq (L^{k(x(t))}V)(x(t)) \\ &= -l(x(t)) - h(k(x(t))). \end{aligned} \tag{A4}$$

Applying the Ito formula [15] to (A4), it follows that $V(x(t))$ is a positive supermartingale, converges almost surely [11] to a random variable $V_\infty$ and, in particular, $\sup_t V(x(t)) < \infty$, almost surely. The Ito formula also yields

$$\begin{aligned} E\left[\int_0^t (l(x(\tau)) + h(k(x(\tau))))\, d\tau\right] &\leq V(x(0)) - E[V(x(t))] \\ &\leq V(x(0)). \end{aligned} \tag{A5}$$

Taking the limit, as $t \to \infty$, we obtain $\int_0^\infty (l(x(\tau)) + h(k(x(\tau))))\, d\tau < \infty$, almost surely. Then invoke Assumption 2 and use Lemma 1 (as in the proof for Case A) to complete the proof.

*Case C (Stochastic Average Cost Problems):* Similarly with (A5) we obtain

$$E\left[\int_0^t l(x(\tau))\, d\tau\right] - gt \leq V(x(0)) - E[V(x(t))] \leq V(x(0)). \tag{A6}$$

Dividing by $t$ and using Fatou's lemma [15] we obtain

$$E\left[\liminf_{t\to\infty} \frac{1}{t}\int_0^t l(x(\tau))\, d\tau\right] \leq \liminf_{t\to\infty} E\left[\frac{1}{t}\int_0^t l(x(\tau))\, d\tau\right] \leq g.$$

Therefore, $\liminf_{t\to\infty}(1/t)\int_0^t l(x(\tau))\,d\tau < \infty$, almost surely, which, in view of Assumption 1c, implies the last part of the theorem. ∎

*Proof of Theorem 2:* The proof of sufficiency is trivial, so we concentrate on the proof of necessity. Since the theorem has to be proved for each $x$ separately, we assume that a particular $x$ has been fixed and we drop the dependence on $x$ from our notation.

Assume, without loss of generality, that there is no $i$ such that $\partial f/\partial u_i(u)$ is identically zero. For any $i$, let $H_i = \{\partial f/\partial u_i(u): u \in R^m\}$. If there exist $u^1, u^2$ such that $\partial f/\partial u_i(u^1)$ and $\partial f/\partial u_i(u^2)$ are not collinear, then there exists a vector $a$ such that $\partial f/\partial u_i(u^1)\cdot a > 0$ and $\partial f/\partial u_i(u^2)\cdot a < 0$, thus contradicting Assumption 3ii). Therefore, $H_i$ is contained in some one-dimensional subspace of $R^m$, which may be represented by some nonzero vector $b_i \in H_i$.

We now partition the set $\{1,\cdots,m\}$ of components of $u$ into a set of classes $A_1,\cdots,A_q$, as follows: two components $i, j$ will belong to the same class if and only if $b_i$ is collinear to $b_j$.

Since $\partial f/\partial u_i$ is collinear to $b_i$, so must be $\partial^2 f/\partial u_i\,\partial u_j$. By interchanging the order of differentiation, we conclude that either $i$ and $j$ belong to the same class, or $\partial^2 f/\partial u_i\,\partial u_j(u) = 0$, $\forall u$. Based on this observation, the representation (17) follows immediately. The fact that the functions $f^k$ must themselves satisfy Assumption 3ii) is also straightforward. ∎

*Proof of Theorem 3:* From any one of the formulas (4), (9), or (11), corresponding to Cases A, B, C, respectively, we obtain, by differentiating with respect to $u$,

$$\frac{\partial h}{\partial u}(k(x)) + \frac{\partial V}{\partial x}(x)\cdot F(x) = 0. \tag{A7}$$

Let $x(t), u(t)$ be the trajectories of the state and the control resulting from the perturbed closed-loop system. Let, for convenience,

$$\begin{aligned} c(t) &= \frac{\partial h}{\partial u}(k(x(t)))[\Phi(k(x(\cdot)))(t) - k(x(t))] \\ &\quad + h(k(x(t))) \\ &= \frac{\partial V}{\partial x}(x(t))F(x(t))[k(x(t)) - \Phi(k(x(\cdot)))(t)] \\ &\quad + h(k(x(t))), \end{aligned} \tag{A8}$$

where the last equality follows from (A7). Note that Assumption 6ii) states that

$$\int_0^t c(\tau)\,d\tau \geq \epsilon\int_0^t h(k(x(\tau)))\,d\tau, \qquad \forall t \geq 0. \tag{A9}$$

*Case A (Deterministic Problems):* Using (4) and (A8)

$$\begin{aligned} \frac{dV}{dt}(x(t)) &= \frac{\partial V}{\partial x}(x(t))\cdot f^0(x(t)) \\ &\quad + \frac{\partial V}{\partial x}(x(t))F(x(t))\Phi(k(x(\cdot)))(t) \\ &= \frac{\partial V}{\partial x}(x(t))\cdot f^0(x(t)) + \frac{\partial V}{\partial x}(x(t))F(x(t))k(x(t)) \\ &\quad + h(k(x(t))) - c(t) \\ &= -l(x(t)) - c(t). \end{aligned} \tag{A10}$$

We now integrate (A10), use (A9) and take the limit as $t \to \infty$ to obtain

$$\int_0^\infty (l(x(\tau)) + \epsilon h(k(x(\tau)))) < \infty$$

and the desired result follows from Lemma 1, as in the proof of Theorem 1.

*Case B (Stochastic Total Cost Problems):* Similarly with (A10), we obtain from (9) and (A8),

$$(L^{u(t)}V)(x(t)) = -l(x(t)) - c(t). \tag{A11}$$

We now integrate (A11) and use the Ito rule to obtain, for $0 \leq t \leq T$,

$$E[V(x(T))|\mathscr{F}_t] - V(x(t)) \leq -E\left[\int_t^T c(\tau)\,d\tau|\mathscr{F}_t\right].$$

Given that $c(t)$ is adapted to $\{\mathscr{F}_t\}$, it follows that

$$E\left[V(x(T)) + \int_0^T c(\tau)\,d\tau|\mathscr{F}_t\right] \leq V(x(t)) + \int_0^t c(\tau)\,d\tau$$

which shows that $V(x(t)) + \int_0^t c(\tau)\,d\tau$ is a (positive) supermartingale and therefore converges. Hence, $V(x(t))$ has bounded sample paths. From (A9), (A11), and the Ito rule we obtain

$$\begin{aligned} &E\left[\int_0^t (l(x(\tau)) + \epsilon h(k(x(\tau))))\,d\tau\right] \\ &\leq E\left[\int_0^t (l(x(\tau)) + c(\tau))\,d\tau\right] \leq V(x(0)), \qquad \forall t \geq 0. \end{aligned} \tag{A12}$$

In view of Assumption 2, $\int_0^\infty (l(x(\tau)) + h(u(\tau)))\,d\tau < \infty$, and the desired result follows from Lemma 1.

*Case C (Stochastic Average Cost Problems):* Similarly with (A12) we obtain

$$E\left[\int_0^t l(x(\tau))\,d\tau\right] \leq V(x_0) + gt, \qquad \forall t \geq 0.$$

This is the same inequality as (A6) in the proof of Theorem 1 and the rest of the proof is the same as in Theorem 1. ∎

*Proof of Proposition 1:* The proof of the first statement is immediate from Parseval's theorem. See also [13, p. 177]. The next statements are [13, Corollaries 4 and 5]. ∎

*Proof of Proposition 2:* It is trivial to check that the integrand in the left-hand side of (18) will be nonnegative for all $\tau$; hence (18) is satisfied. ∎

## References


[1] B. D. O. Anderson and J. B. Moore, *Linear Optimal Control.* Englewood Cliffs, NJ: Prentice-Hall, 1971.

[2] H. W. Bode, *Network Analysis and Feedback Amplifier Design.* New York: Van Nostrand, 1945.

[3] J. C. Doyle and G. Stein, "Multivariable feedback design: Concepts for a classical/modern synthesis," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 4–16, Feb. 1981.

[4] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control.* New York: Springer-Verlag, 1975.

[5] T. Glad, "On the gain margin of nonlinear and optimal regulators," in *Proc. 21st IEEE Conf. on Decision and Contr.*, Orlando, FL, Dec. 1982, pp. 957–962.

[6] I. M. Horowitz, *Synthesis of Feedback Systems.* New York: Academic, 1963.

[7] N. V. Krylov, *Controlled Diffusion Processes.* New York: Springer-Verlag, 1980.

[8] H. J. Kushner, "Optimality conditions for the average cost per unit time problem with a diffusion model," *SIAM J. Contr. Optimiz.*, vol. 16, pp. 330–346, Mar. 1978.

[9] N. A. Lehtomaki, N. R. Sandell, Jr., and M. Athans, "Robustness results in linear-quadratic Gaussian based multivariable control designs," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 75–92, Feb. 1982.

[10] P. L. Lions and J. L. Menaldi, "Optimal control of stochastic integrals and Hamilton–Jacobi–Bellman equations," *SIAM J. Contr. Optimiz.*, vol. 20, pp. 58–81, 82–95, Jan. 1982.

[11] P. A. Meyer, *Probability and Potentials.* Waltham, MA: Blaisdell, 1966.

[12] M. G. Safonov, A. J. Laub, and G. L. Hartmann, "Feedback properties of multivariable systems: The role and use of the return difference matrix," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 47–65, Feb. 1981.
[13] M. G. Safonov and M. Athans, "Gain and phase margin for multiloop LQG regulators," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 173–179, Apr. 1977.
[14] M. G. Safonov, *Stability and Robustness of Multivariable Feedback Systems.* Cambridge, MA: M.I.T. Press, 1980.
[15] E. Wong, *Stochastic Processes in Information and Dynamical Systems.* New York: McGraw-Hill, 1971.
[16] P. K. Wong and M. Athans, "Closed-loop structural stability for linear-quadratic optimal systems," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 94–99, Feb. 1977.

**John N. Tsitsiklis**, for a photograph and biography, see p. 50 of the January 1984 issue of this TRANSACTIONS.

**Michael Athans** (S'58–M'61–SM'69–F'73), for a photograph and biography, see p. 8 of the January 1984 issue of this TRANSACTIONS.

# Optimal Control of a Queueing System with Two Heterogeneous Servers

WOEI LIN, MEMBER, IEEE, AND P. R. KUMAR, MEMBER, IEEE

*Abstract*—**The problem considered is that of optimally controlling a queueing system which consists of a common buffer or queue served by two servers. The arrivals to the buffer are Poisson and the servers are both exponential, but with *different* mean service times. It is shown that the optimal policy which minimizes the mean sojourn time of customers in the system is of threshold type. The faster server should be fed a customer from the buffer whenever it becomes available for service, but the slower server should be utilized if and only if the queue length exceeds a readily computed threshold value.**

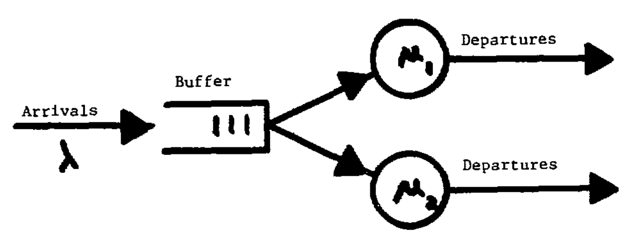



Fig. 1. Queueing system.

## I. INTRODUCTION

THE queueing system shown in Fig. 1 is considered. Arrivals to the buffer form a Poisson process of rate $\lambda$. The buffer is served by two servers with *different* mean service times. The service time of a customer at server $i$ is exponentially distributed with rate parameter $\mu_i$ $(i=1,2)$. Without loss of generality we assume $\mu_1 > \mu_2$. To ensure stability we shall also assume that $\lambda < \mu_1 + \mu_2$. We wish to minimize the mean sojourn time of customers in the queueing system. Note that the sojourn time = waiting time in buffer + service time. By Little's theorem [1], this is equivalent to minimizing the mean number of customers in the system.

Manuscript received January 3, 1983; revised June 15, 1983, July 29, 1983, and August 31, 1983. Paper recommended by A. S. Willsky, Associate Editor for Estimation. This work was supported by the U.S. Army Research Office under Contract DAAG29-80-K0038.

W. Lin was with the Department of Mathematics and Computer Science, University of Maryland Baltimore County, Catonsville, MD 21228. He is now with the Department of Mathematics and Computer Science, Morgan State University, Baltimore, MD.
P. R. Kumar is with the Department of Mathematics and Computer Science, University of Maryland Baltimore County, Catonsville, MD 21228.

If server $i$ is available (i.e., idle) and the buffer is nonempty (i.e., there is a customer waiting for service) should a customer from the buffer be provided to server $i$? We show that the optimal policy governing the dispatching of customers from the buffer to an available server is of *threshold type*, i.e., the faster server, whenever it is available and whenever the buffer is nonempty, should be dispatched a customer, but the slower server should be dispatched a customer only when, at the instant of dispatching, the number of customers in the buffers exceeds a certain readily computed *threshold* value.

This problem, which is a generalization of the $M/M/2$ queue incorporating different service rates at the two servers, was first posed by Larsen [2], who also conjectured that the optimal policy is of threshold type, and proceeded to do a detailed performance analysis of policies of threshold type. The motivation for the queueing system considered here lies in its application to the dynamic routing problem in computer systems or communication networks. For example, what is here called a "server," could be a communication line over which messages can be sent. The "service time" alluded to in this paper is then just the time taken for the message to traverse the line. Messages arriving at the buffer then have to be routed over one of several communication lines, each with a different mean transmission delay, and the goal now is to

0018-9286/84/0800-0696$01.00 