Proceeding of the 2004 American Control Conference
Boston, Massachusetts June 30 - July 2, 2004

Cost Distribution Shaping: The Relations Between Bode Integral, Entropy, Risk-Sensitivity, and Cost Cumulant Control

Chang-Hee Won

***Abstract*— The cost function in stochastic optimal control is viewed as a random variable. Then the classical linear-quadratic-Gaussian control, entropy control, risk-sensitive control, and cost cumulant control can be viewed as the cost distribution shaping methods. In this paper, we will survey the existing relations between entropy, Bode integral, and risk-sensitive cost function. Furthermore, we will relate the cost cumulants with information theoretic entropy, and Bode integral. The interpretation of cost cumulant control is given in terms of the control entropy minimization. The paper also relates information theoretic entropy with exponential-of-integral cost function using a Lagrange multiplier and calculus of variations. Finally, the logarithmic-exponential-of-integral cost function is related to the information theoretic entropy using large deviation theory.**

I. INTRODUCTION

One of the well known results on the fundamental limit on a closed loop system are Bode integral [2, p. 285]. Bode considered frequency response of single input, single output, open-loop stable, linear-time-invariant (LTI) system. He stated that the integral of logarithm of the magnitude of the sensitivity function is equal to zero for an open-loop stable system. An interpretation of the Bode's integral is that reducing the sensitivity due to the system disturbances at one range of frequencies by feedback control will amplify the transients and oscillations at other frequencies. An extension of this Bode's result has been given by Freudenberg and Looze for a general open-loop, multivariable, LTI system [11].

The time varying extension is studied by Iglesias and his coworkers. They studied the relations between the Bode integral, system theoretic entropy, and infinite horizon risk-aversive control cost function [14]. Moreover, Iglesias related Bode integral to the difference of the output and input entropy rates [15]. This result can also be seen from the definition of the entropy rate [20, p. 534]. Consider an LTI system as shown is Figure 1 with input, x_n, output, y_n, the open loop transfer function, $G(z)$, and the sensitivity function, $S(z)$. Now, assume that $S(z)$ is stable. Then the entropy rate is average uncertainty per sample and it is given by

$$\bar{H}(y) = \bar{H}(x) + \frac{1}{2\sigma}\int_{-\sigma}^{\sigma} \log|detS(e^{j\omega})|d\omega. \quad (1)$$

The Bode integral is equal to output entropy rate minus the input entropy rate. The second term in the right hand side is the Bode integral, which is also defined as the system

C. Won is with the Department of Electrical Engineering, University of North Dakota, Grand Forks, ND 58202-7165, U.S.A. won@und.edu

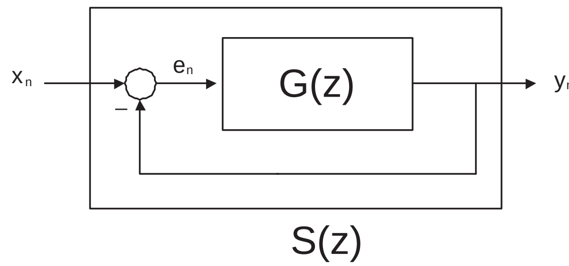

Fig. 1. Feedback System Block Diagram

variety [28], and this is equal to zero if the system is open-loop stable. Moreover, the Bode integral is equal to the sum of open right half plane poles, $\{p_i = 1\}$ if the sytem is open loop unstable. In equations this relationship is given by

$$\int_0^{\pi} \ln|\det S(e^{j\omega})|d\omega = \begin{cases} 0 & \text{if } G(z) \text{ is stable} \\ \sum -\ln|p_i| & \text{if } G(z) \text{ is unstable,} \\ & \text{i.e., } |p_i| < 1 \end{cases}$$

In 1988, Glover and Doyle showed that the system theoretic entropy is related to the infinite horizon risk-aversive cost function [12]. This has been related to cost cumulant control in [23]. The system theoretic entropy is related to the information theoretic entropy through the conditional entropies [21, p. 54].

Also, in 1988 Saridis provided an interpretation of stochastic optimal control in terms of information theoretic entropy [24]. He claimed that the cost mean optimization is equivalent to control entropy minimization where the entropy density function is found to be the worst possible case. Figure 2 summarizes the relationship between various control methods. The researchers who related the two areas are given near the arrows. Figure 2 also shows the section number of this paper where the relations between two areas are established.

Consider the cost function and the control action in stochastic optimal control as a random variable. Then we can view various control problems as the cost distribution shaping methods. Moments or cumulants characterize a distribution, thus by optimizing a particular moment such as the mean, we are shaping the distribution of the cost function. Conventionally, in optimal stochastic control, one establishes a cost function and optimizes the controller with respect to the mean of the cost function. Optimizing only the mean (the first cumulant) of the cost function is a special case of optimizing the distribution of the cost function. For instance, the cost variance (second cumulant)

0-7803-8335-4/04/$17.00 ©2004 AACC

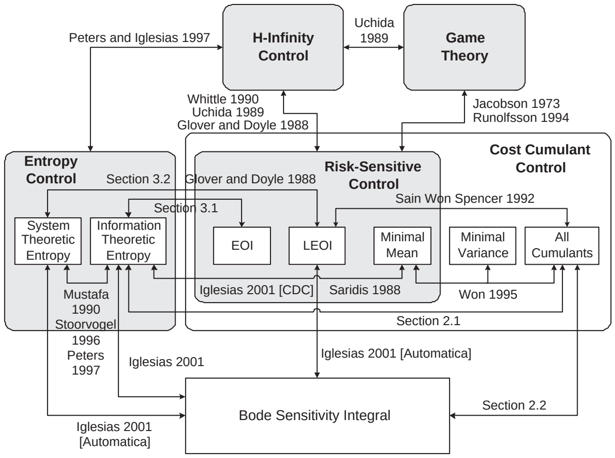

Fig. 2. Relations Among Various Robust Control Methods

can be minimized. The variance indicates to what extent performance is spread around its mean. In some instances this variance is an important parameter to optimize. For example, if a manufacturer wants to produce products above a certain quality level, the desirable production quality distribution will be a sharp distribution with small variance with the mean pushed close to the rejection threshold. Figure 3 shows how the distributions change as the first three cumulants vary. This can be achieved if one can control any cumulant of the cost function.

The information theoretic entropy is related to the sum of all the cost cumulants or moments, thus entropy optimization is also a form of cost distribution shaping method. The risk-sensitive cost function is also an infinite sum of all the cumulants or moments[27], thus it is also a form of cost distribution shaping method.

Cost moment or cumulant control can also be interpreted in terms of entropy minimization, which is discussed in Section II-A. Also cost cumulant control is related to the Bode integral in Section II-B. We relate the entropy with the risk-sensitive cost function in Section III. Because risk-sensitive control can be formulated in terms of exponential-of-integral (EOI) or logarithm-exponential-of-integral (LEOI) cost functions, we will relate these two cases with information theoretic entropy. Then the conclusions are given in the final section.

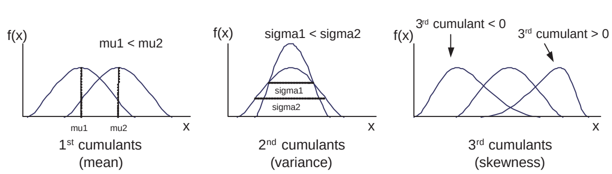

Fig. 3. Effects of Cumulants on the Distribution

II. COST CUMULANT CONTROL, ENTROPY, AND BODE INTEGRAL

In this section we relate cost cumulant control with information theoretic entropy, and Bode integral.

A. Cost Cumulant Control and Information Theoretic Entropy

Cost cumulant control is a form of cost distribution shaping method, and it is closely related to entropy control. We will show that the optimal control that minimization the entropy of the control is equivalent to minimizing n-th moment of the cost function. Because moments are related to cumulants [18], cost cumulant control is related to minimum entropy control. In this paradigm, the cost function represents the energy of a system and optimization is performed over the admissible control laws.

Define entropy as $h = -\int p(u) \ln p(u) dx$ where $p(u)$ is the density function. In minimum entropy control, we find the density that would give the maximum entropy. Then we determine the control law, u, that would minimize the entropy when the density, $p(u)$ is the worst case density. Conceptually this is similar to H_∞ control, where the infinity norm of the transfer function is minimized. Consequently, entropy control and H_∞ control are related [12].

We are ready to formulate the cost cumulant control in terms of information theoretic entropy. Consider a nonlinear system,

$$\dot{x} = f(x, u, w, t). \tag{2}$$

and a nonquadratic cost function,

$$\hat{J} = \int_0^{t_F} c(x, u, t) dt. \tag{3}$$

Then the n-th moment cost function is given as

$$J = E\{\hat{J}^n\}, \tag{4}$$

where $n = 1, 2, \cdots$. If we consider $\hat{J}$ as a random variable then we can optimize this cost function in many different ways. We could minimize any of the moments or cumulants of the cost function. Moments and cumulants characterize a distribution. Thus, we could optimize the cost distribution to suit our control purposes.

Now, we will provide an interpretation of cost cumulant control in terms of entropy. Let $p(u)$ be the probability density of the control action and let Ω_u denote a set of admissible controls. Then

$$\int_{\Omega_u} p(u) dx = 1. \tag{5}$$

For $p(u)$ we may assign the following entropy,

$$H(u, p) = -\int_{\Omega_u} p(u) \ln p(u) dx = -E_{\Omega_u}\{\ln p(u)\}. \tag{6}$$

We would like to minimize the entropy with the constraints (4) and (5). So we create an unconstrained expression for the entropy using Lagrange multipliers.

$$\begin{aligned} I &= H(u) - \mu[E\{\hat{J}^n\} - J] - \lambda_1 \left[\int_{\Omega_u} p(u)dx - 1\right] \\ &= -\int_{\Omega_u} p(u)\ln p(u)dx - \mu \int_{\Omega_u} p(u)\hat{J}^n dx + \mu J \\ &\quad -\lambda_1 \int_{\Omega_u} p(u)dx + \lambda_1. \end{aligned}$$

Using calculus of variations (for exampe, see [3, p. 47]), we maximize I with respect to $p(u)$,

$$\frac{\partial}{\partial p}\left[-p\ln p - \mu p\hat{J}^n - \lambda_1 p\right] = 0$$
$$-\ln p - 1 - \mu\hat{J}^n - \lambda_1 = 0.$$

Therefore, the worst case density is

$$p(u) = \exp(-\lambda - \mu\hat{J}^n),$$

where $\lambda = \lambda_1 + 1$. And the corresponding entropy is

$$H(u) = \lambda + \mu E\{\hat{J}^n\}. \tag{7}$$

Theorem 2.1: Consider the following system

$$\begin{aligned} \frac{dx}{dt} &= f(x,u,t) + g(x,t)w(t); \quad x(t_0) = x_0, \\ z(t) &= h(x,t) + v(t) \end{aligned}$$

where w and v are independent identically distributed (i.i.d.) random processes. A necessary and sufficient condition for $u^*(x,t)$ to minimize $J = E\{\hat{J}^n\}$ subject to the above dynamic constraints is that $u(x,t)$ minimizes the entropy $H(u(x,t))$ where the associated $p(u(x,t))$ is the maximum entropy density function satisfying Jaynes's maximum entropy principle.

Proof: The proof will closely follow the minimum mean case of Saridis [24]. For the necessary condition, consider Jayne's maximum entropy condition

$$H(x_0,u,p) = \lambda + \mu E\{\hat{J}^n(x_0,u,t)\}$$

where μ is the constant satisfying

$$\int_{\Omega_u} p(u)dx = 1 \text{ and } \frac{\partial H}{\partial p} = 0.$$

Then

$$\begin{aligned} \min_u H(x_0,u,p) \quad &\Leftrightarrow \quad \frac{\partial H}{\partial u} = 0 \\ &\Leftrightarrow \quad \min_u E\{\hat{J}^n(x_0,u,t)\} \\ &\qquad = \int_{\Omega_u} \min \hat{J}^n p dx. \end{aligned}$$

Using the lemma of the calculus of variations this is equivalent to minimizing $\hat{J}^n p$ or

$$\frac{\partial \hat{J}^n}{\partial u} = \frac{\partial \hat{J}^n}{\partial u}p + \hat{J}^n\frac{\partial p}{\partial u} = 0. \tag{8}$$

But $\frac{\partial p}{\partial u} = -\mu\frac{\partial \hat{J}^n}{\partial u}p$ and from Eq. (8) we obtain

$$(1 - \mu\hat{J}^n)\frac{\partial \hat{J}^n}{\partial u}p = 0.$$

For $p \neq 0$ and $\mu \neq 1$, it implies that $\frac{\partial \hat{J}^n}{\partial u} = 0$ or $u^* = \hat{J}^n(x_0,u^*,t)$ is minimum.

For sufficient condition we assume that the conditions of interchanging integrations and minimizations are satisfied. Then u^* that minimizes $\hat{J}^n(x_0,u)$ implies that it maximizes $p(x_0,u)$ such that

$$\begin{aligned} H(u^*) &= -\int_{\Omega_u} p(u^*)\ln p(u^*)dx \\ &= \lambda + \int_{\Omega_u} p(u^*)\mu\hat{J}^n dx \\ &= \lambda + \int_{\Omega_u} \max_u p(u) \min_u \mu\hat{J}^n dx \\ &= \lambda + \int_{\Omega_u} \min_u p(u) \min_u \mu\hat{J}^n dx \\ &\Rightarrow \lambda + \int_{\Omega_u} \min_u [p(u)\mu\hat{J}^n]dx \\ &= \min_u H(u) \end{aligned}$$

where $\max_u(\cdot) = -\min_u(\cdot)$ and $\min_u(A)\min_u(B) \Rightarrow \min_u(AB)$ for $A, B > 0$. $\square$

Now, we consider a few special cases. If f is linear, L is quadratic and w is Gaussian process. Moreover, if $p(u(x,t))$ is the worst case (maximum entropy) density function and $E\{\hat{J}\}$ is the cost function. Then look for u that minimizes the above cost function such that u minimizes the entropy $H(u) = -\int_{\Omega_u} p(u)\ln p(u)dx$ where $p(u(x,t))$ is the maximum entropy density function satisfying Jayne's maximum entropy principle. This is Saridis's interpretation of the stochastic optimal control in terms of information theoretic entropy. Saridis determined the worst case density to be in exponential form. Here it is interesting to note that at least from 1955 it was known that for a single variable case, the distribution that gives the maximum entropy for a given mean and variance is a Gaussian distribution [1, p. 257]. Another interesting fact is that the conditional entropy is equal to the entropy rate for white and Markov processes.

B. *Cost Cumulant Control and Bode Integral*

This section relates the cost cumulant cost function with Bode integral. We will relate the risk-sensitive cost funciton to the cost cumulant cost function as in [23]. Then we will relate risk-sensitive cost function to the Bode integral following the steps in [14]. This will result in relating the cost cumulant cost function with the Bode integral.

Consider a linear time-invariant (LTI) system, $e_k = Sx_k$, as shown in Figure 1. The corresponding quadratic cost function is

$$J_D = \sum_{k=0}^{T-1} x_k' x_k - e_k' e_k,$$

where T is the terminal time. The risk-sensitive cost function is given as

$$J_{RS}(J_D, \theta) = -\frac{2}{\theta T} \log E \exp\left(-\frac{\theta}{2} J_D\right). \tag{9}$$

The first charateristic function of J_D is defined as

$$\phi(s) = E\{\exp(-sJ_D)\} = 1 + \sum_{i=1}^{\infty} \frac{(-1)^i}{i!} \alpha_i s^i, \tag{10}$$

where α_i's are the i-th moments of J_D, and the second characteristic function is given by

$$\psi(s) = \log \phi(s) = \sum_{i=1}^{\infty} \frac{(-1)^i}{i!} \beta_i s^i, \tag{11}$$

where β_i's are i-th cumulants of J_D. From Eqs. (10) and (11), we obtain

$$\log E\{\exp(-sJ_D)\} = \sum_{i=1}^{\infty} \frac{(-1)^i}{i!} \beta_i s^i. \tag{12}$$

Then substitute Eq. (12) into (9).

$$J_{RS}(J_D, \theta) = -\frac{2}{\theta T} \sum_{i=1}^{\infty} \frac{(-1)^i}{i!} \beta_i \left(\frac{\theta}{2}\right)^i. \tag{13}$$

The above equation gives the relationship between the risk-sensitive cost function and the cost cumulant cost function.

Following [14], we will now derive the relationship between the risk-sensitive cost function and the Bode integral. Assume the sensitivity function, $S(z)$, satisfies $\|S\|_\infty < \alpha$. Then by spectral factorization, we obtain

$$I - \alpha^{-2} S^*(z) S(z) = G^*(z) G(z)$$

and

$$\log |det S(e^{j\omega})| = \frac{1}{2} \log det\left(I - G^*(e^{j\omega}) G(e^{j\omega})\right) + m \log \alpha,$$

where m is the matrix dimension of the sensitivity function. This can be rewritten as

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \log |det S(e^{j\omega})| d\omega = m \log \alpha + \frac{1}{4\pi} \int_{-\pi} \pi \log det\left(I - G^*(e^{j\omega}) G(e^{j\omega})\right) d\omega. \tag{14}$$

From [12], we have

$$\lim_{T\to\infty} \left(-\frac{2}{\theta T} \log E \exp\left(-\frac{\theta}{2} J_D\right)\right) = = \frac{1}{2\pi\theta} \int_{-\pi}^{\pi} \log det\left(I - G^*(e^{j\omega}) G(e^{j\omega})\right) d\omega, \tag{15}$$

if $\|G\|_\infty^2 < 1/|\theta|$. We assume $\theta = -1$, which is the risk-aversive case, and substitute Eq. (15) into (14) to obtain

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \log |det S(e^{j\omega})| d\omega = m \log \alpha - \lim_{T\to\infty} \left(\frac{1}{T} \log E \exp\left(\frac{1}{2} J_D\right)\right). \tag{16}$$

From Eq. (9), we obtain

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \log |det S(e^{j\omega})| d\omega = m \log \alpha - \lim_{T\to\infty} \frac{1}{T} \sum_{i=1}^{\infty} \frac{1}{i!} \left(\frac{1}{2}\right)^i \beta_i. \tag{17}$$

We have used Eqs. (9) and (13) on (16) to obtain the last equality. Eq. (17) relates the Bode integral with the cost cumulants of J_D.

For LTI and infinite horizon case, the above equation provides an interpretation of the Bode integral in terms of the cost cumulants. The Bode integral is zero for an open loop stable system, and it is a sum of unstable poles for an open loop unstable system. Thus, a sum of linear combination of all the cumulants are equal to $m \log \alpha$ for the open loop stable system, and $m \log \alpha$ plus a sum of unstable poles for the stable system.

Also if we substitute the Bode integral in Eq. (17) to (1), we note that the entropy rate of the ouput minus the input is the sum of all cumulants. This relates system variety to the cumulants, and provides an interpretation of conservation of the sum of cost cumulants.

III. ENTROPY AND RISK-SENSITIVE COST FUNCTION

Risk-sensitive control can be viewed as optimizing the infinite sum of all the moments or the cumulants of the cost function. The moment case is related to the exponential-of-integral (EOI) cost function and the cumulant case to the logarithm-exponential-of-integral (LEOI) cost function [27]. EOI control is also known as linear exponential quadratic Gaussian (LEQG) control. In the next subsection we will relate the information theoretic entropy to the EOI cost function. Subsequently, we will relate entropy with the LEOI cost function.

A. Entropy and EOI Cost Function

Here, a form of risk-sensitive cost function, EOI cost function, is related to the information theoretic entropy. We consider the nonlinear system given in Eq. (13), and a risk-sensitive cost function,

$$J = E\{\exp(\theta \hat{J})\}, \tag{18}$$

where $\hat{J}$ is given in (3). Now, we consider the density function for the controller:

$$p(u) = p(u(x,t)),$$

and

$$\int_{\Omega_u} p(u)dx = 1.$$

Then the information theoretic entropy is given as

$$H(u) = -\int_{\Omega_u} p(u)\ln p(u)dx = -E_{\Omega_u}\{\ln p(u)\}.$$

To find the density function that would maximize the entropy, we form an unconstrained expression for the entropy using Lagrange multipliers.

$$\begin{aligned} I &= H(u) - \mu[E\{\exp(\theta\hat{J})\} - J] - \lambda_1[\int_{\Omega_u} p(u)dx - 1] \\ &= -\int_{\Omega_u} p(u)\ln p(u)dx - \mu\int_{\Omega_u} p(u)\exp(\theta\hat{J})dx + \mu J \\ &\quad -\lambda_1\int_{\Omega_u} p(u)dx + \lambda_1. \end{aligned}$$

Using calculus of variations, we maximize I with respect to $p(u)$,

$$\frac{\partial}{\partial p}[-p\ln p - \mu p\exp(\theta\hat{J}) - \lambda_1 p] = 0$$

$$-\ln p - 1 - \mu exp(\theta\hat{J}) - \lambda_1 = 0.$$

Therefore, the worst case density is determined as

$$p(u) = \exp(-\lambda - \mu\exp(\theta\hat{J})),$$

where $\lambda = \lambda_1 + 1$. And the corresponding entropy is

$$H(u) = \lambda + \mu E\{\exp(\theta\hat{J})\}. \tag{19}$$

The interpretation in terms of EOI cost function is as follows. The minimization of EOI cost function (18) is equivalent to minimizing the information theoretic entropy given by Eq. (19). Thus, in EOI control we are finding the controller that would minimize the entropy (19) assuming the worst case density function, $p(u)$.

B. Entropy and LEOI Cost Function

More generally, in this Section we relate information theoretic entropy (Shannon's entropy) to the LEOI cost function using large deviation theory [4], [7], [8], [10]. We still consider the general nonlinear state equation given by Eq. (2). As in [22], we define the occupation distribution of x_t in Γ as

$$\Pi_{t_F}(\Gamma) = \frac{1}{t_F}\int_0^{t_F} \chi_\Gamma(x_t)dt,$$

where χ_Γ is the indicator of Γ, $\Gamma \subset \mathbb{R}^n$, and $t \in [0, t_F]$. The infinite horizon LEOI cost function is given as

$$\lambda(j) = \lim_{t_F\to\infty}\frac{1}{t_F}\log E\left\{\exp\left(\int_0^{t_F} j(x_t)\,dt\right)\right\}. \tag{20}$$

where $j : \mathbb{R}^n \to \mathbb{R}$ is a continuous function. Let μ be finite measure on $\mathbb{R}^n$, and define for $f \in C_b(\mathbb{R}^n)$ (bounded continuous space on $\mathbb{R}^n$)

$$\langle\langle\mu, f\rangle\rangle = \int_{\mathbb{R}^n} f(x)\mu(dx).$$

(20) can be rewritten using the occupation distribution as

$$\lambda(j) = \lim_{t_F\to\infty}\frac{1}{t_F}\log E\left\{\exp\left(t_F\langle\langle\Pi_{t_f}, j\rangle\rangle\right)\right\}. \tag{21}$$

In Donsker and Varadhan's notation [6], we note that $\Pi_t(\Gamma) = L_{t,w}(A)$. The asymptotic rate involving functionals of $L_{t,w}$ is then governed by the I-function (see [5, page 390])

$$I(\mu) = -\inf_{u\in\mathcal{D}^+}\int_X\left(\frac{Lu}{u}\right)(x)\mu(dx), \tag{22}$$

where $\mathcal{D}^+$ is the set of positive functions u in the domain of L. It was shown by Donsker [6] that this I-function is related to the entropy function $H(Q)$ by

$$I(\mu) = \inf_{\substack{Q\in\mathcal{M}_S(\Omega)\\ q(Q)=\mu}} H(Q),$$

where Ω be a space of functions $\omega(\cdot)$ on $-\infty < t < \infty$ with values in a Polish space X. $\mathcal{M}_S(\Omega)$ denotes the space of stationary processes on Ω, and $q(Q) = \mu$ means the marginal of the stationary measure Q is μ. We want to relate this entropy function to the more familiar information theoretic entropy (Shannon's entropy) of the form

$$h(\lambda;\mu) = -\int f(x)\log f(x)\lambda(dx)$$

if $h(\lambda;\mu)$ is finite. Let X, Σ be a measurable space and λ, μ be the probability measures on (X, Σ). For each $\omega \in \Omega$ we denote by $\omega(t)$ the value of the function $\omega(\cdot)$ at time t. We also denote by Ω_t^+, the corresponding space of functions on $[t, \infty)$ with values in X and we denote by $\mathcal{F}_t^s$ the σ-algebra in Ω generated by $\omega(\sigma)$ for $s \le \sigma \le t$.

Note that $h(\lambda;\mu)$ is finite if and only if (a) μ is absolutely continuous with respect to λ, and (b) the Radon-Nikodym derivative, $\frac{d\mu}{d\lambda} = f(x)$, is such that $f(x)\log f(x) \in L_1(\lambda)$. Let $\{P_{t,x}\}$ be a homogeneous Markov family of measures on Ω_t^+, define $\{P_{t,\omega(t)}\} = P_{t,x}$ with the starting point $x = \omega(t)$, and define $\{Q_{t,\omega}\}$ as the regular conditional probability distributions of Q given $\mathcal{F}_t^{-\infty}$. Finally, we have the relationship

$$H(Q) = H(1, Q) = -E\left\{h_{\mathcal{F}_1^0}(P_{0,\omega(0)}; Q_{0,\omega})\right\},$$

where $\mathcal{F}_t^0$ is the σ-fields in Ω generated by $\omega(\sigma)$ for $0 \le \sigma \le t$. Thus $H(Q)$ is the entropy of the stationary process Q with respect to the Markov process $P_{0,x}$ at time 1. In the infinite horizon LEOI control problem, we are minimizing (20) which corresponds to minimizing (21). To minimize (21), we should minimize the exponent, which implies that we are minimizing the occupation distribution $L_{t,\omega}(A)$. Furthermore, from (22), minimizing $L_{t,\omega}(A)$ corresponds to maximizing $I(\mu)$, and maximizing $I(\mu)$ corresponds to maximizing the entropy $H(Q)$. And finally, maximizing $H(Q)$ corresponds to minimizing $h(\lambda;\mu)$. Thus, as in the EOI case the LEOI optimization is a minimum entropy control method.

IV. CONCLUSIONS

In this paper, we surveyed various relations between entorpy control, risk-sensitive control, cost cumulant control, and Bode integral. Then we provided an interpretation of cost moment/cumulant control in terms of cost distribution shaping and entorpy minimization. The infinite sum of all the cumulants is related to the entropy rate and system variety. Finally, we related risk-sensitive cost function with the information theoretic entropy. Thus providing an interpretation for risk-sensitive control as an entropy minimization method.

REFERENCES

[1] M. S. Bartlett, *An Introduction to Stochastic Processes with Special Reference to Methods and Applications,* Third Edition, Cambridge University Press, Cambridge, 1978.

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

[3] A. E. Bryson, Jr., and Y.-C. Ho, *Applied and Optimal Control, Optimization, Estimation, and Control,* Revised Printing, New York: Hemisphere Publishing Co., 1975.

[4] J.-D. Deuschel and D. W. Stroock, *Large Deviations*, San Diego: Academic Press, 1989.

[5] M. D. Donsker and R. S. Varadhan, "Asymptotic Evaluation of Certain Markov Process Expectations for Large Time, III," *Communications on Pure and Applied Mathematics*, Vol. XXIX, pp. 389–461, 1976.

[6] M. D. Donsker and R. S. Varadhan, "Asymptotic Evaluation of Certain Markov Process Expectations for Large Time, IV," *Communications on Pure and Applied Mathematics*, Vol. 36, pp. 183–212, 1983.

[7] P. Dupuis, and R. S. Ellis, *A Weak Convergence Approach to the Theory of Large Deviations,* Wiley-Interscience, 1997.

[8] R. S. Ellis, *Entropy, Large Deviations, and Statistical Mechanics,* New York: Springer-Verlag, 1985.

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

[10] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems,* New York: Springer-Verlag, 1984.

[11] J. S. Freudenberg and D. P. Looze, "Right Half Plane Poles and Zeros and Design Tradeoffs in Feedback Systems," *IEEE Transactions on Automatic Control,* Vol. AC-30, No. 6, pp. 555–565, June 1985.

[12] K. Glover and J. C. Doyle, "State-Space Formulae for All Stabilizing Controllers that Satisfy an H_∞-Norm Bound and Relations to Risk Sensitivity," *Systems and Control Letters*, Volume 11, pp. 167–172, 1988.

[13] K. Glover, "Minimum Entropy and Risk-Sensitive Control: The Continuous Time Case," *Proceedings 28th IEEE Conference on Decision and Control*, pp. 388–391, December 1989.

[14] P. A. Iglesias, "Tradeoffs in linear time-varying systems: an analogue of Bode's sensitivity integral," *Automatica,* Vol. 37, pp. 1541–1550, 2001.

[15] P. A. Iglesias, "An analogue of Bode's integral for stable non linear systems: Relations to entropy," *Proceedings of the 40th IEEE Conference on Decision and Control,* Orlando, FL, USA, pp. 3419–3420, December 2001.

[16] D. H. Jacobson, "Optimal Stochastic Linear Systems with Exponential Performance Criteria and Their Relationship to Deterministic Differential Games," *IEEE Transactions on Automatic Control*, Volume AC-18, pp. 124–131, 1973.

[17] E. T. Jaynes, "Information Theory and Statistical Mechanics," *Phys. Rev.*, Vol. 106 1957.

[18] A. Stuart and J. K. Ord, *Kendall's Advanced Theory of Statistics*, Volume 1 "Distribution Theory," Fifth Edition, Oxford University Press, New York, 1987.

[19] *Minimum entropy H_∞ control,* Vol. 146 of Lecture Notes in Control and Information Sciences, Springer Verlag, Berlin, 1990.

[20] A. Papoulis, *Probability, Random Variables, and Stochastic Processes,* Second Edition, McGraw-Hill Book Company, New York, 1984.

[21] M. A. Peters and P. A. Iglesias, *Minimum Entropy Control for Time-Varying Systems,* Systems and Control: Foundations & Applications, Birkhauser, Boston, 1997.

[22] T. Runolfsson, "The Equivalence between Infinite-Horizon Optimal Control of Stochastic Systems with Exponential-of-Integral Performance Index and Stochastic Differential Games," *IEEE Transactions on Automatic Control*, Vol. 39, No. 8, pp. 1551–1563, 1994.

[23] M. K. Sain, C.-H. Won, and B. F. Spencer, Jr. "Cumulant Minimization and Robust Control," *Stochastic Theory and Adaptive Control, Lecture Notes in Control and Information Sciences 184*, T. E. Duncan and B. Pasik-Duncan (Eds.), Springer-Verlag, pp. 411–425, 1992.

[24] G. Saridis, "Entropy Formulation of Optimal and Adaptive Control," *IEEE Transactions on Automatic Control*, Vol. 33, No. 8, pp. 713–721, 1988.

[25] A. A. Stoorvogel and J. H. Van Schuppen, "System Identification with Information Theoretic Criteria," *Identification, Adaptation, Learning: The Science of Learning Models from Data (NATO Asi Series. Series F, Computer and Systems Sciences, Vol. 153*, Sergio Bittanti (Editor), Giorgio Picci (Editor), Springer, Berlin, 1996, pp. 289-338.

[26] K. Uchida and M. Fujita, "On the central controller: Characterizations via differential games and LEQG control problems," *Systems & Control Letters*, Volume 13, pp9–13, 1989.

[27] C.-H. Won, "Cost Cumulants in Risk-Sensitive and Minimal Cost Variance Control," Ph. D. Dissertation, Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN, USA, May 1995.

[28] Hui Zhang and Youxian Sun, "Bode Integral and Laws of Variety in Linear Control System," *Proceedings of the American Control Conference,* Denver, CO, pp. 66–70, June 4-6, 2003.