model-based controllers that robustly stabilize all plants in the uncertainty region. This measure therefore gives us guidelines to select the uncertainty region that is best tuned for robust stability analysis among all available ones. To illustrate the impact of our results in terms of the connection between identification and robust control, we return to the example above. With our robust stability measure for uncertainty sets, we were able to conclude that the $G_{\rm mod}$-based controller set that is guaranteed to robustly stabilize $\mathcal{D}^{CL}$ is much larger than the set that is guaranteed to robustly stabilize $\mathcal{D}^{OL}$. Hence, in terms of identification for control, the closed-loop identification design that led to the uncertainty set $\mathcal{D}^{CL}$ is a much better experiment design than the open-loop design that led to $\mathcal{D}^{OL}$. The results of this paper have thus allowed us to establish a connection between identification design and stability robustness of the controllers resulting from such design.

References

[1] X. Bombois, M. Gevers, and G. Scorletti. Controller validation based on an identified model. presented at CDC'99 Tech. Rep. 99.06, CESAME.

[2] ——, "Controller validation for a validated model set," in *Proc. European Control Conf. CD-ROM*, Karlsruhe, Germany, 1999, paper 869.

[3] X. Bombois, M. Gevers, G. Scorletti, and B. D. O. Anderson, "Controller validation for stability and performance based on an uncertainty region designed from an identified model," in *Proc. IFAC Symp. System Identification CD-ROM*, Santa Barbara, CA, 2000, paper 058.

[4] B. Boulet and B. A. Francis, "Consistency of open-loop experimental frequency-response data with coprime factor plant models," *IEEE Trans. Automat. Contr.*, vol. 43, pp. 1680–1691, Dec. 1998.

[5] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, "Linear matrix inequalities in systems and control theory," in *Studies in Appl. Math., vol. 15*. Philadelphia, PA: SIAM, June 1994.

[6] J. Chen, "Frequency-domain tests for validation of linear fractional uncertain models," *IEEE Trans. Automat. Contr.*, vol. 42, pp. 748–760, June 1997.

[7] R. G. Hakvoort, "System identification for robust process control," Ph.D. dissertation, Delft Univ. Technol., Delft, The Netherlands, 1994.

[8] R. L. Kosut, "Uncertainty model unfalsification: A system identification paradigm compatible with robust control design," in *Proc. Conf. Decision and Control*, New Orleans, LA, 1995.

[9] L. Ljung, *System Identification: Theory for the User* , 2nd ed. Upper Saddle, NJ: Prentice-Hall, 1999.

[10] M. Milanese, "Learning models from data: The set membership approach," in *Proc. American Control Conf.*, Philadelphia, PA, 1998, pp. 178–182.

[11] K. Poolla, P. Khargonekar, A. Tikku, J. Krause, and K. Nagpal, "A time-domain approach to model validation," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 951–959, May 1994.

[12] M. G. Safonov and T.-C. Tsao, "The unfalsified control concept and learning," *IEEE Trans. Automat. Contr.*, vol. 42, pp. 843–847, June 1997.

[13] G. Vinnicombe, "Frequency domain uncertainty and the graph topology," *IEEE Trans. Automat. Contr.*, vol. 38, pp. 1371–1383, 1993.

[14] ——, "Measuring the robustness of feedback systems," Ph.D. dissertation, Cambridge Univ., 1993.

Random Spherical Uncertainty in Estimation and Robustness

B. T. Polyak and P. S. Shcherbakov

***Abstract*—A theorem is formulated that gives an exact probability distribution for a linear function of a random vector uniformly distributed over a ball in n-dimensional space. This mathematical result is illustrated via applications to a number of important problems of estimation and robustness under spherical uncertainty. These include parameter estimation, characterization of attainability sets of dynamical systems, and robust stability of affine polynomial families.**

***Index Terms*—Estimation, random noise, robust stability, uncertain systems.**

I. Introduction

Traditionally, different fields of control theory exploit various models for the uncertainty. For instance, in parameter and state estimation, the standard approach deals with random (specifically, Gaussian) perturbations, and least squares and Kalman filtering are the most popular tools for estimation under such assumptions. Later, the model of unknown-but-bounded perturbations was developed, which led to ellipsoidal techniques for estimation [14], [8], [10].

On the other hand, the models of parametric uncertainty in control theory are basically deterministic, e.g., see [1], [2], and [5] devoted to robust stability and performance of uncertain linear systems. One of the drawbacks of such models is that the admissible ranges for the uncertainty that satisfy performance specifications are calculated against the worst case uncertainty, which may happen very rarely in practice. Also, the computational complexity of the methods often grows exponentially in the dimension of the uncertainty vector.

In practical applications, it is quite often the case that hard bounds on the uncertainty are not known. Instead, certain probabilistic characteristics for the uncertain parameters are available, the conclusions are obtained in the form of confidence estimates, and the solution often involves Monte Carlo simulations; see [13], [9], [4], and [7]. Along with low computational complexity, the main benefit is a considerable enhancement of admissible uncertainty domains in exchange of a small probability risk that the deterministic specifications are violated. The results obtained so far relate to independent random variables.

Following the probabilistic approach, in this paper we work with an important class of dependent random parametric uncertainty, namely, with the uniform distribution on a ball in l_2-norm. There are several reasons for such an uncertainty model. First, if the uncertainty is supposed to be of a stochastic nature, the l_2-constraint is associated with a bound on the total energy of random noise; in that case, the random l_2 uncertainty can be thought of as a bridge between probabilistic models and unknown-but-bounded models (with ellipsoidal models of uncertainty as conventional tools), e.g., see Section IV, where the deterministic result is enhanced via its probabilistic counterpart. On the other hand, for the parametric uncertainty, the l_2 model is quite natural, since often, the information about the uncertain parameters is derived from

Manuscript received May 24, 1999; revised February 10, 2000. Recommended by Asociate Editor, R. Tempo. This work was supported by the Russian Foundation for Basic Research under Grant 99-01-00340, by the National Science Foundation under Grant ECS-9418709, and by INTAS under Grant 97-10782.

The authors are with the Institute for Control Science, Moscow 117806, Russia.

Publisher Item Identifier S 0018-9286(00)10002-9.

a preliminary estimation or identification process; the results are usually available in the form of confidence ellipsoids. Finally, we mention the paper [3], where spherically uniform distributions were studied in the context of distributional robustness. Namely, under certain conditions, a (truncated) spherically uniform distribution comes into play as a result of the theory rather than as an ad hoc assumption; it is this distribution that should be adopted in order to obtain reliable conclusions about the performance of the system using the Monte Carlo methods.

II. A Linear Function in a Uniform Distribution on a Ball

The following notation will be used throughout the paper: $(\cdot,\,\cdot)$ is the inner product in $\mathbb{R}^n$ and $\|\cdot\|$ is the Euclidean norm $\|x\|^2 = \sum_{i=1}^n x_i^2$; $\mathbf{B} \doteq \{x \in \mathbb{R}^n\colon \|x\| \le 1\}$ denotes the unit ball in $\mathbb{R}^n$ centered at the origin and $\partial\mathbf{B} \doteq \{x \in \mathbb{R}^n\colon \|x\| = 1\}$ is its surface; $\mathcal{U}(\mathbf{Q})$ is the uniform distribution with support set $\mathbf{Q} \subset \mathbb{R}^n$; $\mathcal{N}(0,\, I_m)$ is the m-dimensional Gaussian distribution with zero mean and identity covariance matrix; $\mathcal{B}(\nu_1,\,\nu_2)$ is the beta distribution with $(\nu_1,\,\nu_2)$ degrees of freedom [16]; $\mathcal{C}(m)$ is the χ^2 distribution with m degrees of freedom; $\Gamma(\cdot)$ is the gamma-function; the symbol $\sim$ means "distributed as"; and E is the symbol of the mathematical expectation.

A. Characterization of the Uniform Distribution

Our main concern in this paper is a random vector having uniform distribution on a ball in $\mathbb{R}^n$, and we present a technical lemma that provides a proper characterization of this distribution.

Lemma 1: Let $\xi \sim \mathcal{N}(0,\, I_n)$. Then $\xi/\|\xi\| \sim \mathcal{U}(\partial\mathbf{B})$. If, in addition, a random variable $\rho \sim \mathcal{U}([0,\,1])$ is independent of ξ, then

$$\rho^{1/n}\frac{\xi}{\|\xi\|} \sim \mathcal{U}(\mathbf{B}).$$

This result is known from the literature, e.g., the first proposition of the lemma can be traced back to [11]; for a more detailed exposition, where more general l_p-norms are used, see [6]. The convenience of the representation provided by Lemma 1 is that we can use Gaussian random variables when generating vectors uniformly distributed on a ball.

The following theorem constitutes the mathematical background of the results in this paper.

Theorem 1: Let a random vector $q \in \mathbb{R}^n$ be uniformly distributed on the unit ball $\mathbf{B} \subset \mathbb{R}^n$. Assume that a matrix $A \in \mathbb{R}^{m\times n}$ has rank $m \le n$. Then the random variable

$$\tau \doteq \left(\left(AA^T\right)^{-1}Aq,\, Aq\right)$$

has the beta distribution $\tau \sim \mathcal{B}(m/2,\,(n-m/2)+1)$ with density

$$f_\tau(x) = \begin{cases} \dfrac{\Gamma\left(\dfrac{n}{2}+1\right)}{\Gamma\left(\dfrac{m}{2}\right)\Gamma\left(\dfrac{n-m}{2}+1\right)}\, x^{(m/2)-1}(1-x)^{(n-m/2)}, \\ \qquad \text{for } 0 < x < 1 \\ 0, \qquad \text{otherwise.}\end{cases} \tag{1}$$

Proof: By construction, the matrix $B \doteq A^T(AA^T)^{-1}A$ has m eigenvalues equal to 1 and $n-m$ eigenvalues equal to 0. Indeed, since B is symmetric, it has n orthonormal eigenvectors e_i, and since $\mathrm{rank}(A) = m$, exactly m of these eigenvectors have the property $Ae_i \ne 0$. Writing $A^T(AA^T)^{-1}Ae_i = \lambda_i e_i$ and premultiplying both sides by A, we obtain $Ae_i = \lambda_i Ae_i$. Then, for $Ae_i \ne 0$, we have $\lambda_i = 1$ with multiplicity m. Otherwise, if $Ae_i = 0$, then $\lambda_i e_i = 0$, whence $\lambda_i = 0$ with multiplicity $n-m$. Therefore, the representation $B = PG_mP^T$ is valid with some orthogonal matrix P, $P^T = P^{-1}$, and

$$G_m = \left(\begin{array}{c|c} I_m & 0_{m,\,n-m} \\ \hline 0_{n-m,\,m} & 0_{n-m,\,n-m}\end{array}\right)$$

where I_m and $0_{\cdot,\,\cdot}$ denote the identity matrix and zero matrices of appropriate dimensions. Then the distribution of $\tau = (Bq,\,q)$ is the same as that of $(G_mq,\,q)$, since the vector $y = P^Tq$ is also distributed uniformly on the ball $\mathbf{B}$. We now introduce the random variables $x_i = q_i^2$, $i = 1,\,\cdots,\,n$, and note that the vector $x = (x_1,\,\cdots,\,x_n)$ has the Dirichlet distribution $\mathcal{D}(1/2,\,\cdots,\,1/2;1)$ with density

$$\frac{\Gamma\left(\dfrac{n}{2}+1\right)}{\left(\Gamma\left(\dfrac{1}{2}\right)\right)^n}\, x_1^{-1/2}\cdots x_n^{-1/2}. \tag{2}$$

See [16, (7.7.1)]. Indeed, under the change of variables $q \to x$, each of the intersections of the unit ball in the q-space with any of 2^n orthants transforms into the unit simplex in the x-space (since $x_i \ge 0$, $i = 1,\,\cdots,\,n$, and $\sum_{i=1}^n x_i \le 1$). The Jacobian of this transformation is equal to $(1/2)^n x_1^{-1/2}\cdots x_n^{-1/2}$ in each of the 2^n intersections so that for the density of x, we have

$$\begin{aligned} f(x_1,\,\cdots,\,x_n) &= 2^n\,\frac{1}{\mathbf{Vol}(\mathbf{B})}\,\frac{1}{2^n}\,x_1^{-1/2}\cdots x_n^{-1/2} \\ &= \frac{1}{\mathbf{Vol}(\mathbf{B})}\,x_1^{-1/2}\cdots x_n^{-1/2}\end{aligned}$$

where $1/\mathbf{Vol}(\mathbf{B})$ is the density of q. This coincides with the density of the Dirichlet distribution $\mathcal{D}(1/2,\,\cdots,\,1/2;1)$ (2), since $\mathbf{Vol}(\mathbf{B}) = (\pi^{n/2}/\Gamma((n/2)+1)) = ((\Gamma(1/2))^n/\Gamma(n/2+1))$. Finally, from, [16, 7.7.2 and 7.7.4] it follows that the random variable $(G_mq,\,q) \doteq \sum_{i=1}^m q_i^2 = \sum_{i=1}^m x_i$ has the beta distribution $\mathcal{B}(\nu_1+\cdots+\nu_m;\, \nu_{m+1}+\cdots+\nu_{n+1}) = \mathcal{B}((m/2),\,((n-m)/2)+1)$. □

Remark 1—Linear Transformation of a Ball: The image of a ball under linear transformation is an ellipsoid, i.e., for $A \in \mathbb{R}^{m\times n}$, $\mathrm{rank}(A) = m \le n$, the set $\{x = Aq\colon q \in \gamma\mathbf{B} \subset \mathbb{R}^n\}$ is the ellipsoid $\mathbf{E} \doteq \{x \in \mathbb{R}^m\colon x^T(AA^T)^{-1}x \le \gamma^2\}$. Theorem 1 can be thought of as a probabilistic counterpart of this well-known fact in the sense that it provides an explicit description of a confidence ellipsoid for a random vector that is a linear transformation of a uniform distribution on a ball. This ellipsoid is defined by the matrix AA^T, and its size is given by the quantile of the beta distribution.

Remark 2—Independence of A: The only conditions imposed on the matrix $A \in \mathbb{R}^{m\times n}$ in Theorem 1 are that it has full rank and $m = \mathrm{rank}(A) \le \dim q = n$. That is, for all such matrices, the resulting distribution of τ is the same as that for $A^0 \doteq (I_m|0_{m,\,n-m})$, and it depends only on $\mathrm{rank}(A)$ and $\dim q$.

Remark 3—Asymptotic Characterization: The result in Theorem 1 can be formulated in the following asymptotic form.

Theorem 2: Assume that for every $n \ge m$, a matrix $A_n \in \mathbb{R}^{m\times n}$ has rank m, and $q^{(n)}$ denotes a random vector uniformly distributed on the unit ball in $\mathbb{R}^n$. Then the random vector $\eta^{(n)} \doteq n^{1/2}(A_nA_n^T)^{-1/2}A_nq^{(n)}$ tends in distribution to $\mathcal{N}(0,\,I_m)$ as $n \to \infty$.

Proof: By Lemma 1, $q^{(n)} = \rho^{1/n}\xi^{(n)}/\|\xi^{(n)}\|$, where $\xi^{(n)} \doteq (\xi_1,\,\cdots,\,\xi_n) \sim \mathcal{N}(0,\,I_n)$. We have $A_n\xi^{(n)} \sim \mathcal{N}(0,\,A_nA_n^T)$, therefore, $(A_nA_n^T)^{-1/2}A_n\xi^{(n)} \sim \mathcal{N}(0,\,I_m)$. Hence, for the random vector η_n, we have the representation $\eta_n = \rho^{1/n}\eta/\zeta_n$, where $\eta \sim \mathcal{N}(0,\,I_m)$ and $\zeta_n \doteq \sqrt{\xi_1^2+\cdots+\xi_n^2}/\sqrt{n}$. Next, $\rho^{1/n} \to 1$ in probability, and by the weak law of large numbers ([16, Theorem 4.3.1] with $x_i = \xi_i - 1$), we obtain $\zeta_n \to 1$ in probability. Therefore, by [16, Theorem 4.3.6a] [relations (4.3.19) and (4.3.20)], we have $\eta_n \to \eta$ in probability, and [16, Theorem 4.3.4] yields $\eta_n \to \eta$ in distribution. □

Since $\|\xi\|^2 \sim \mathcal{C}(m)$ for $\xi \sim \mathcal{N}(0, I_m)$, we arrive at a corollary: $n((A_n A_n^T)^{-1} A_n q_n, A_n q_n) \to \mathcal{C}(m)$ in distribution. Therefore, an important conclusion drawn from Theorem 1 is that a linear transformation changes the nature of the uniform distribution on a ball. Namely, with increase of $\dim q$, the rank of the transformation being unaltered [or, equivalently, with decrease of $\operatorname{rank}(A)$, $\dim q$ being unaltered], the distribution of the transformed vector Aq tends to concentrate closer to the center of the image of the support set rather than to its surface. It is this effect that will be exploited in the sections to follow when constructing *probabilistic predictors* of certain sets in $\mathbb{R}^n$.

Remark 4—The Case $m = 2$: In control applications, the case $m = 2$ is important because of the complex plane being the range. See Section V; i.e., $\tau \sim \mathcal{B}(1, n/2)$. Then from (1), we have

$$f_\tau(x) = f_{1,\,n/2}(x) = \begin{cases} \dfrac{n}{2}(1-x)^{(n/2)-1}, & \text{for } 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

and we obtain the explicit expression for the cumulative distibution function (cdf)

$$F_\tau(x) = F_{1,\,n/2}(x) = \begin{cases} 0, & \text{for } x < 0 \\ 1-(1-x)^{n/2}, & \text{for } 0 \le x \le 1 \\ 1, & \text{for } x > 1. \end{cases}$$

Remark 5—Ellipsoidal Support: A formally more general case of the uniform distribution $q \sim \mathcal{U}(\mathbf{E})$ over the ellipsoid $\mathbf{E} = \{x \in \mathbb{R}^n\colon x^T M^{-1} x \le 1 \text{ for some } M > 0\}$ reduces to the spherical setting by introducing the scaled matrix $\tilde{A} = AM^{-1/2}$.

III. Parameter Estimation in Linear Regression

In this section, Theorem 1 is applied to the problems of parameter estimation. First, as an illustrative example, we consider the simplest problem of estimating a vector parameter from the measurements corrupted by spherically dependent noise.

Let $c^* \in \mathbb{R}^m$ be an unknown vector and $y_i = c^* + \xi_i$, $i = 1, \cdots, n$, $n \ge m$, be its measurements corrupted by noise ξ_i such that the composite vector $\xi = (\xi_1 \cdots \xi_n)$ is uniformly distributed on the mn-dimensional ball of radius r: $\xi \sim \mathcal{U}(r\mathbf{B}) \subset \mathbb{R}^{mn}$. Then, by straightforward manipulations (or using Theorem 1) it can be shown that the arithmetic mean $\hat{c}_n = \sum_{i=1}^n y_i/n$ is an unbiased linear estimate of c^*, and its accuracy is given by

$$\mathrm{E}\|\hat{c}_n - c^*\|^2 = \frac{r^2 m}{mn^2 + 2n}.$$

Hence, for spherically uniform noise, the variance of the estimate above is of order $O(1/n^2)$. Note that there is no contradiction between $O(1/n^2)$ and the standard estimate with independent noise, where the variance is $O(1/n)$. The reason is that in the spherical setting, the level of noise in each observation decreases as the number of observations grows, since the total energy of noise is bounded from above.

Next, we consider the linear regression model

$$y_i = a_i^T c^* + \xi_i, \qquad i = 1, \cdots, n \tag{3}$$

where $c^* \in \mathbb{R}^m$, $m \le n$ is the vector of unknown parameters, $a_i \in \mathbb{R}^m$, $i = 1, \cdots, n$ are fixed known regressors, y_i, $i = 1, \cdots, n$ are observations, and $\xi_i \in \mathbb{R}$, $i = 1, \cdots, n$ is noise such that the vector $\xi \doteq (\xi_1, \cdots, \xi_n)$ has the uniform distribution on the ball $r\mathbf{B} \subset \mathbb{R}^n$ of radius $r > 0$. Assuming that there are m linearly independent vectors among the a_i, the least squares estimate for c^* is

$$\hat{c} = \left(A^T A\right)^{-1} A^T y$$

where $y \doteq (y_1, \cdots, y_n)^T$ and $A = [a_1^T\, a_2^T \cdots a_n^T] \in \mathbb{R}^{n \times m}$ is the regression matrix composed of a_i. The following theorem provides a closed-form description of confidence ellipsoids for the unknown parameter.

Theorem 3: Let $0 \le p \le 1$, and let τ_p denote the $100p$% quantile of the beta distribution $\mathcal{B}(m/2, (n-m/2)+1)$. Then, under the conditions above, the ellipsoid

$$\mathbf{E}_p \doteq \left\{x \in \mathbb{R}^k\colon \left(A^T A(x-\hat{c}),\, x-\hat{c}\right) \le r^2 \tau_p\right\} \tag{4}$$

is a $100p$% confidence domain for the vector c^* in (3).

Proof: We have $\hat{c} = c^* + B\xi$, where $B \doteq (A^T A)^{-1} A^T$. We introduce the random variable $\tilde{\xi} \doteq \xi/r$ uniformly distributed on the unit ball and consider the residual $\hat{c} - c^* = rB\tilde{\xi}$. By Theorem 1, the random variable

$$\tau \doteq \left(\left(r^2 BB^T\right)^{-1}(\hat{c} - c^*),\, \hat{c} - c^*\right) \tag{5}$$

is distributed $\mathcal{B}(m/2, (n-m/2)+1)$, and by the definition of the $100p$% quantile, we have

$$\operatorname{Prob}\left\{\left(\left(r^2 BB^T\right)^{-1}(\hat{c} - c^*),\, \hat{c} - c^*\right) \le \tau_p\right\} = p.$$

Noting that $(BB^T)^{-1} = A^T A$, this relation writes $\operatorname{Prob}\{(A^T A(\hat{c} - c^*), \hat{c} - c^*) \le r^2 \tau_p\} = p$, or, equivalently, $\operatorname{Prob}\{c^* \in \mathbf{E}_p\} = p$, where $\mathbf{E}_p$ is the ellipsoid given by (4); it is defined by its matrix $(A^T A)^{-1}$ and has center $\hat{c}$ and size $r\tau_p^{1/2}$. □

Remark 6: For $m = 2$, the explicit expression for the cdf is available (see Remark 4 in the previous section), and we have

$$\tau_p = 1 - (1-p)^{2/n}. \tag{6}$$

To compute τ_p in the general case, one can use statistical tables for quantiles of the beta distribution (e.g., see [12]).

IV. Attainability Sets of Dynamical Systems

In this section, we study a discrete-time dynamical system described by the equation

$$x_{k+1} = Ax_k + Bw_{k+1}, \qquad k = 0, 1, \cdots,$$
$$A \in \mathbb{R}^{n \times n},\ B \in \mathbb{R}^{n \times m},\ w_k \in \mathbb{R}^m \tag{7}$$

where the uncertainty $\{w_1, \cdots, w_N\}$ accumulated by the Nth step satisfies the constraints

$$\sum_{k=1}^N \|w_k\|^2 \le c^2 \qquad \text{for some } c > 0. \tag{8}$$

By definition, for every $N > 0$, the *attainability set* of system (7), (8) is the set of all its possible states that can be attained by the Nth step [14]; this is formally written as

$$\mathbf{Q}_N = \left\{x_N = A^N x_0 + A^{N-1} Bw_1 + A^{N-2} Bw_2 + \cdots + ABw_{N-1} + Bw_N\colon \sum_{k=1}^N \|w_k\|^2 \le c^2\right\}.$$

Said another way, the set $\mathbf{Q}_N$ is generated by the uncertainty w in the system description and characterizes the resulting uncertainty in the state, accumulated by the Nth step. For a controllable pair (A, B), this set is the ellipsoid

$$\mathbf{Q}_N = \mathbf{E}N \doteq \left\{x \in \mathbb{R}^n\colon \left(S_N^{-1}(x - \overline{x}_N),\, x - \overline{x}_N\right) \le c^2\right\}$$

with center

$$\overline{x}_N \doteq A^N x_0 \tag{9}$$

and matrix $S_N = M_N M_N^T$, where

$$M_N \doteq \begin{bmatrix} A^{N-1}B & A^{N-2}B & \cdots & AB & B \end{bmatrix} \in \mathbb{R}^{n \times mN} \tag{10}$$

is the controllability matrix of system (7). For the case of Schur stable matrix A (i.e., $\max_k |\lambda_k(A)| < 1$), there exists a limiting attainability ellipsoid

$$\mathbf{E}_\infty \doteq \lim_{N\to\infty} \mathbf{E}N = \left\{x \in \mathbb{R}^n\colon \left(S_\infty^{-1}x,\, x\right) \le c^2\right\}$$

where $S_\infty > 0$ is a solution of the discrete matrix Lyapunov equation $AS_\infty A^T + BB^T = S_\infty$ (indeed, S_N satisfies the recursive relation $S_{N+1} = AS_N A^T + BB^T$, and for a stable matrix A we arrive at $S_\infty = AS_\infty A^T + BB^T$ in the limit as $N \to \infty$).

Assume now that the uncertainty is random; specifically, the mN-dimensional random vector $w \doteq (w_1;\ w_2;\ \cdots;\ w_N)$ has a uniform distribution on the unit ball $\mathbf{B} \subset \mathbb{R}^{mN}$; write $w \sim \mathcal{U}(\mathbf{B})$. We represent

$$x_N = \overline{x}_N + cM_N w$$

where M_N and $\overline{x}_N$ are defined by (10) and (9). As in Section III, consider the random variable $\tau_N = ((c^2 M_N M_N^T)^{-1}(x_N - \overline{x}_N),\, x_N - \overline{x}_N)$, for which Theorem 1 yields $\tau_N \sim \mathcal{B}(n/2,\, (mN - n/2) + 1)$. Specifying the confidence probability p and denoting the $100p$% quantile for the beta distribution by τ_{N_p}, we arrive at the following theorem.

Theorem 4: The ellipsoid

$$\mathbf{E}_{N,\,p} = \left\{x \in \mathbb{R}^n\colon \left(S_N^{-1}(x - \overline{x}_N),\, x - \overline{x}_N\right) \le c^2 \tau_{N_p}\right\}$$

is a $100p$% predictor of the attainability set $\mathbf{E}_N$ in the sense that $\mathrm{Prob}\{x_N \in \mathbf{E}_{N,\,p}\} = p$.

With increase of N ($n,\, m,\, p$ being fixed), the quantile $\tau_{N,\,p}$ decreases, and Theorem 4 gives a considerable *volumetric reduction*

$$\frac{\mathbf{Vol}(\mathbf{E}_N)}{\mathbf{Vol}(\mathbf{E}_{N,\,p})} = 1/\tau_{N,\,p}^n.$$

For instance, with $n = 2$, $m = 1$, $p = 0.99$, and $N = 30$, we obtain $\tau_N p \approx 0.2644$ so that at the thirtieth step, the reduction constitutes $1/\tau_{N,\,p}^2 \approx 14.3092$. Note that the reduction does not depend on A and B, but only on the specification N and p and the dimensions m and n.

This is interpreted to mean that using a probabilistic description allows for the high-confidence replacement of the state uncertainty $\mathbf{E}_N$ with the much smaller uncertainty set $\mathbf{E}_{N,\,p}$. Moreover, if there exists a limiting attainability set $\mathbf{E}_\infty$, from Theorem 2 and the subsequent discussion, it follows that as $N \to \infty$, the limiting distribution of x_N is a singular distribution on $\mathbf{E}_\infty$ so that in the limit, $\mathbf{E}_{N,\,p}$ shrinks to the point $x = 0$.

V. Robust Stability of Spherical-Affine Families of Polynominals

In this section, a probabilistic criterion of robust stability is developed for affine polynomial families under spherical uncertainty using the result in Theorem 1.

A. Spherical Polynomial Families and Deterministic Criteria

We consider the following *spherical-affine* polynomial family:

$$p(s,\, q) = p_0(s) + \sum_{i=1}^{n} q_i p_i(s) \tag{11}$$

where $p_0(s)$ is the nominal and $p_i(s)$, $i = 1, \cdots, n$, are known perturbation polynomials with real coefficients, $q = (q_1, \cdots, q_n)^T$ is the vector of uncertain parameters subjected to spherical constraints

$$q \in \gamma\mathbf{B} \doteq \left\{q \in \mathbb{R}^n\colon \sum_{i=1}^{n} q_i^2 \le \gamma^2\right\} \tag{12}$$

and $\gamma > 0$ is the range of uncertainties. We consider the continuous-time case and Hurwitz stability. The nominal polynomial $p_0(s)$ is assumed to be stable and there is no degree dropping, i.e., $p(s,\, q)$ have the same degree for all $q \in \gamma\mathbf{B}$. The goal is to check if (11), (12) is robustly stable and to determine the robust stability radius

$$\gamma_{\max} \doteq \max\{\gamma\colon p(s,\, q) \text{ is stable for all } q \in \gamma\mathbf{B}\}.$$

The traditional approach to checking the robust stability is via describing the value set of the family and applying the zero exclusion principle [2]. By definition, the *value set* of family (11), (12) is the set $\mathbf{V}(\omega) = \{p(j\omega,\, q)\colon q \in \gamma\mathbf{B}\} \subset \mathbb{C}$. In the sequel, complex numbers z are identified with two-dimensional vectors $(\mathrm{Re}\, z,\, \mathrm{Im}\, z)^T$, and sets on the complex plane are considered as sets in $\mathbb{R}^2$. Introducing the matrix

$$A = A(\omega) = \begin{bmatrix} \mathrm{Re}\, p_1(j\omega) & \mathrm{Re}\, p_2(j\omega) & \cdots & \mathrm{Re}\, p_n(j\omega) \\ \mathrm{Im}\, p_1(j\omega) & \mathrm{Im}\, p_2(j\omega) & \cdots & \mathrm{Im}\, p_n(j\omega) \end{bmatrix}$$

we represent $p(j\omega,\, q) = p_0(j\omega) + A(\omega)q$; then the value set of (11) and (12) is the ellipse

$$\mathbf{E}(\omega) = \left\{x \in \mathbb{R}^2\colon \left(\left(A(\omega)A^T(\omega)\right)^{-1}(x - p_0(j\omega)),\; x - p_0(j\omega)\right) \le \gamma^2\right\}. \tag{13}$$

See Remark 1 of Section II. By the zero exclusion principle, the deterministic robust stability radius $\gamma_{\max}$ for family (11), (12) is determined as

$$\gamma_{\max} = \sup_{\omega \ge 0}\{\gamma\colon 0 \notin \mathbf{E}(\omega)\}$$

that is, by constructing the value set and performing the frequency sweep. Keeping (13) in mind, we obtain

$$\gamma_{\max} = \inf_{\omega \ge 0} \varphi(\omega)$$
$$\varphi(\omega) = \left(\left(A(\omega)A^T(\omega)\right)^{-1} p_0(j\omega),\, p_0(j\omega)\right)^{1/2}. \tag{14}$$

Results of this kind were obtained in [15].

B. The Probabilistic Approach

Assume that the specified uncertainty radius γ is greater than the robustness margin $\gamma_{\max}$ obtained via a deterministic test of the sort (14). Then, adopting the probabilistic viewpoint and assuming that the uncertainty vector q is random, uniformly distributed over $\gamma\mathbf{B}$, a natural direct estimate for the *probability of stability* of (11) and (12) might be $\tilde{P} \doteq \mathbf{Vol}(\gamma_{\max}\mathbf{B})/\mathbf{Vol}(\gamma\mathbf{B}) = (\gamma_{\max}/\gamma)^n$. It is seen that for high dimensions n, this quantity can be very small even for the values of γ, that are only slightly greater than $\gamma_{\max}$. Below, using the result of Theorem 1, we provide a much finer estimate for the probability of stability.

First, we introduce the notion of *probabilistic predictor* of the value set (13). For any fixed $\omega \ge 0$, a domain $\mathbf{E}_{1-\varepsilon}(\omega) \subset \mathbf{E}(\omega)$ is called a $100\%(1-\varepsilon)$-predictor of $\mathbf{E}(\omega)$ if $\mathrm{Prob}\{p(j\omega,\, q) \in \mathbf{E}_{1-\varepsilon}(\omega)\} = 1-\varepsilon$, where $\varepsilon \in [0,\, 1]$ is some prespecified *probability risk*. Using the result of Theorem 1, it can be easily shown that

$$\mathbf{E}_{1-\varepsilon}(\omega) \doteq \left\{x \in \mathbb{R}^2\colon \left(\left(A(\omega)A^T(\omega)\right)^{-1}(x - p_0(j\omega)),\; x - p_0(j\omega)\right) \le \gamma_\varepsilon^2\right\}$$
$$\gamma_\varepsilon = \gamma\sqrt{1 - \varepsilon^{2/n}}.$$

By construction, the value set $\mathbf{E}(\omega)$ and its predictor $\mathbf{E}_{1-\varepsilon}(\omega)$ are similar ellipses centered $p_0(j\omega)$, and the quantity $\gamma_\varepsilon/\gamma = (1-\varepsilon^{2/n})^{1/2} < 1$ is the similarity coefficient. The reciprocal quantity

$$\alpha(\varepsilon) \doteq \left(1 - \varepsilon^{2/n}\right)^{-1/2} > 1$$

is called the *risk-adjusted enhancement* of the uncertainty domain in the sense that if the uncertainty vector q is chosen randomly in the *enhanced* ball $\alpha(\varepsilon)\gamma\mathbf{B}$, then the associated random point $p(j\omega, q)$ belongs to the true value set with probability $1 - \varepsilon$. We stress that in contrast to the results in the probabilistic robustness literature obtained so far, an analytic expression for the probabilistic enhancement is made possible.

Now, in order to establish the robust stability *with probability* ε and to obtain the probabilistic stability margin, it would be reasonable to perform the frequency sweep with the predictor $\mathbf{E}_{1-\varepsilon}(\omega)$ instead of the true value set $\mathbf{E}(\omega)$. Note, however, that the construction above relates to any fixed value of frequency; therefore, to make any conclusions about the probability of stability, similar estimates are to be made uniformly over a range of frequencies. The first attempt to overcome this so-called cross-frequency effect is due to Chen and Zhou [7]. We propose a different way of solving the problem, which is based on the lemma below. Note that we deal with the $n \times n$ ω-dependent matrix $B(\omega) \doteq A^T(\omega)(A(\omega)A^T(\omega))^{-1}A(\omega)$ defined over a finite range Ω of frequencies.

Lemma 2: Assume that a symmetric matrix $B(\omega) \in \mathbb{R}^{n\times n}$ has rank 2 and depends on the parameter $\omega \in \Omega \doteq [\omega_1, \omega_2]$, and let $\omega_0 \in \Omega$. Assume next that $\|B(\omega) - B(\omega_0)\| \le \delta$ for all $\omega \in \Omega$ and some small $\delta > 0$. Then for $q \sim U(\mathbf{B})$ and any $\varepsilon \in [0, 1]$, we have

$$\begin{aligned} \mathrm{Prob}\left\{(B(\omega)q, q) \le 1 - \varepsilon^{2/n} \quad \forall\, \omega \in \Omega\right\} \\ \ge 1 - \left(\varepsilon^{2/n} + \delta\right)^{n/2}. \end{aligned} \tag{15}$$

Proof: According to Theorem 1, the random variable $(B(\omega_0)q, q)$ is distributed via $\mathcal{B}(1, n/2)$ so that for a prespecified $\varepsilon \in [0, 1]$, we have

$$\mathrm{Prob}\left\{(B(\omega_0)q, q) \le 1 - \varepsilon^{2/n}\right\} = 1 - \varepsilon$$

and, respectively

$$\mathrm{Prob}\left\{(B(\omega_0)q, q) \le 1 - \varepsilon^{2/n} - \delta\right\} = 1 - (\varepsilon^{2/n} + \delta)^{n/2}.$$

See Remark 4 in Section II for the explicit form of the cdf for $\mathcal{B}(1, n/2)$. From the condition $\|B(\omega) - B(\omega_0)\| \le \delta$ for all $\omega \in \Omega$, it follows that

$$(B(\omega)q, q) \le (B(\omega_0)q, q) + \delta \qquad \text{for all } \omega \in \Omega$$

and the three relations above yield the assertion of the lemma. □

This lemma is interpreted to mean the following. Let the classical robustness margin $\gamma_{\max}$ be determined via (14), and let γ be greater than $\gamma_{\max}$ so that $\varphi(\omega) \le \gamma$ for $\omega_1 \le \omega \le \omega_2$. We write $\Omega \doteq [\omega_1, \omega_2]$ and refer to Ω as the violation segment (the range of frequencies for which the zero exclusion principle is violated for a given γ). Then, by picking a certain middle point ω_0 in the violation segment and computing the maximum variation δ of the matrix $B(\omega)$ over Ω, we estimate the relative volume of the violating portion of the γ-ball $\gamma\mathbf{B}$ in the q-space rather than computing the difference between $\mathbf{Vol}(\gamma\mathbf{B})$ and $\mathbf{Vol}(\gamma_{\max}\mathbf{B})$. The example below demonstrates how Lemma 2 can be applied to the estimation of the probability of stability.

Example: The experiment was conducted with the continuous-time polynomial family $p(s) = a_1 + a_2 s + \cdots + a_n s^{n-1}$ with real coefficients subjected to ellipsoidal constraints of the form $\sum_{i=1}^{n}((a_i - a_i^0)/\beta_i)^2 \le \gamma^2$, where $a^0 \doteq (a_1^0, \cdots, a_n^0)$ are the nominal values of the coefficients and $\beta \doteq (\beta_1, \cdots, \beta_n)$ are known shaping factors. Note that the representation above reduces to the form of (11) and (12) by introducing the nominal polynomial $p_0(s) \doteq a_1^0 + a_2^0 s + \cdots + a_n^0 s^{n-1}$, perturbation polynomials $p_i(s) \doteq \beta_i s^{i-1}$, $i = 1, \cdots, n$, and uncertainties $q_i \doteq (a_i - a_i^0)/\beta_i$, $i = 1, \cdots, n$, bounded to the ball $\gamma\mathbf{B}$. Specifically, in the experiment we work with

$$\begin{aligned} a^0 = [&195.0 \quad 1162.0 \quad 3056.0 \quad 4950.0 \quad 5553.0 \quad 4505.0 \\ &2701.0 \quad 1200.0 \quad 388.0 \quad 88.0 \quad 13.0 \quad 1.0] \end{aligned}$$

and

$$\begin{aligned} \beta = [&19.0 \quad 116.0 \quad 305.0 \quad 495.0 \quad 555.0 \quad 450.0 \\ &270.0 \quad 120.0 \quad 38.8 \quad 8.8 \quad 1.3 \quad 0.1] \end{aligned}$$

i.e., $\deg p(s) = 11$, there are $n = 12$ uncertain parameters, and the nominal is Hurwitz stable, $\max_k \mathrm{Re} s_k = -0.4948$; s_k denotes the roots of $p_0(s)$.

For this family, the deterministic criterion (14) gives $\gamma_{\max} \approx 0.4584$, which is attained at $\omega_0 \approx 1.68$. Assume next that the uncertainty vector q is bounded to the ball of radius $\gamma = 1.1\gamma_{\max} \approx 0.5042$, which defines the violation segment $\Omega = [\omega_1, \omega_2]$ with $\omega_1 \approx 1.36$ and $\omega_2 \approx 2.08$. We take ω_0 to be the middle point in the violation segment and compute numerically the value of δ as $\delta \doteq \max_{\omega\in\Omega} \|B(\omega) - B(\omega_0)\|$, where $B(\omega) = A^T(\omega)(A(\omega)A^T(\omega))^{-1}A(\omega)$ and shown at the equation at the bottom of the page. This results in $\delta \approx 0.2657$. By (15), we find the estimated probability of stability approximately equal to 0.8489, while the volumetric estimate gives the unsatisfactory result $\tilde{P} = (1/1.1)^{12} \approx 0.3186$.

It should be noted that (15) in Lemma 2 gives a lower bound of the probability of stability, and the "true" probability is higher. Indeed, if at some frequency, the trajectory of a system driven by a particular uncertainty $q \in \gamma\mathbf{B}$ does not belong to the value set, then this does not imply its instability. In other words, the "probabilistic" violation of the zero exclusion principle does not necessarily yield the loss of robust stability. To illustrate, we performed a direct stability test, which consists in straightforward sampling the uncertainty in the γ-ball $\gamma\mathbf{B}$ and counting up the number of stable polynomials thus generated. For 100 000 sampling values of $q \in \gamma\mathbf{B}$, we detect 99 990 stable polynomials, i.e., the "true" probability of stability is estimated to be 0.9999.

VI. Conclusion

In this paper, we studied the behavior of systems subjected to the class of random uncertainties uniformly distributed on a ball. The mathematical result of Theorem 1 is illustrated via a number of applications; the approach exposes advantages over classical techniques.

The results admit generalizations along a number of directions. It is felt that many problems in systems theory can be solved for the uncertainty model studied in this paper: in filtering, this will lead to an analog of Kalman filtering under spherical-uniform noise; in design,

$$A(\omega) = \begin{bmatrix} \beta_1 & 0 & -\beta_3\omega^2 & 0 & \beta_5\omega^4 & 0 & -\beta_7\omega^6 & 0 & \beta_9\omega^8 & 0 & -\beta_{11}\omega^{10} & 0 \\ 0 & \beta_2\omega & 0 & -\beta_4\omega^3 & 0 & \beta_6\omega^5 & 0 & -\beta_8\omega^7 & 0 & \beta_{10}\omega^9 & 0 & -\beta_{12}\omega^{11} \end{bmatrix}$$

controllers can be obtained to attenuate external perturbations which are bounded to a ball and have a uniform distribution. Similar results are expected in robust control, maximum likelihood parameter estimation, etc.

ACKNOWLEDGMENT

The authors would like to express their gratitude to the anonymous referees, whose comments and suggestions contributed much to the quality of this paper. In particular, a more direct and simpler proof of Theorem 1 was hinted at by one of the referees.

REFERENCES

[1] J. Ackermann, *Robust Control: Systems with Uncertain Physical Parameters*. New York: Springer-Verlag, 1993.
[2] B. R. Barmish, *New Tools for Robustness of Linear Systems*. New York: Macmillan, 1994.
[3] B. R. Barmish, C. M. Lagoa, and R. Tempo, "Radially truncated uniform distribution for probabilistic robustness of control systems," in *Proc. Amer. Control Conf.*, Albuquerque, NM, 1997.
[4] B. R. Barmish and B. T. Polyak, "A new approach to open robustness problems based on probabilistic prediction formulae," in *Proc. 13th IFAC World Congress*, vol. H, San Francisco, CA, July 1996, pp. 1–6.
[5] S. P. Bhattacharyya, H. Chapellat, and L. H. Keel, *Robust Control: The Parametric Approach*. Upper Saddle River, NJ: Prentice-Hall, 1995.
[6] G. Calafiore, F. Dabbene, and R. Tempo, "Uniform sample generation of vectors in l_p balls for probabilistic robustness analysis," in *Proc. 37th CDC*, Tampa, FL, 1998.
[7] X. Chen and K. Zhou, "'On the probabilistic characterization of model uncertainty and robustness," in *Proc. 36th CDC*, vol. 4, San Diego, CA, 1997, pp. 3816–3825.
[8] F. L. Chernous'ko, *State Estimation for Dynamic Systems*. Boca Raton, FL: CRC Press, 1994.
[9] X. Zhu, Y. Huang, and J. Doyle, "Soft vs. hard bounds in probabilistic robustness analysis," in *Proc. 35th CDC*, Kobe, Japan, 1996.
[10] A. B. Kurzhanskii and I. Valyi, *Ellipsoidal Calculus for Estimation and Control*. Basel: Birkhäuser, 1997.
[11] M. E. Muller, "A note on a method for generating random points uniformly on n-dimensional spheres," *Commun. ACM*, vol. 2, pp. 19–20, 1959.
[12] K. Pearson, *Tables of the Incomplete Beta Function*. Cambridge, U.K.: Cambridge Univ. Press, 1934.
[13] L. R. Ray and R. F. Stengel, "A Monte Carlo approach to the analysis of control systems robustness," *Automatica*, vol. 29, no. 1, pp. 229–236, 1993.
[14] F. C. Schweppe, *Uncertain Dynamic Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1973.
[15] C. B. Soh, C. S. Berger, and K. P. Dabke, "On the stability properties of polynomials with perturbed coefficients," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 1033–1036, Oct. 1985.
[16] S. Wilks, *Mathematical Statistics*. New York: Wiley, 1963.

Stability Analysis and Bang–Bang Sliding Control of a Class of Single-Input Bilinear Systems

Yon-Ping Chen, Jeang-Lin Chang, and Kuo-Ming Lai

***Abstract*—This paper introduces a novel bang–bang sliding control of a class of single-input bilinear systems. The sliding function is chosen via the well-known pole-assignment method for linear time-invariant systems. Importantly, the bang–bang sliding control generates a reaching-and-sliding region and a stable-sliding region, each expressed by a set of linear inequalities. Both regions comprise the equilibrium point, shown to be asymptotically stable. However, the stability analysis is processed under the limitation that the system state should be initially located in the reaching-and-sliding region. Two numeric examples are used for demonstration.**

***Index Terms*—Bang–bang sliding control, bilinear systems, stability analysis.**

I. INTRODUCTION

In general, bilinear systems are expressed by a state differential equation, which is linear in control and linear in state but not jointly linear in state and control. Bilinear systems have been found in diverse processes and fields (see [1] and [2] for an excellent introduction) and many control strategies have developed, such as the quadratic feedback control [3], [4] and the sliding-mode control [5], [6]. Here, we will focus on a class of "bang–bang" single-input bilinear systems with controller designed by the sliding-mode theory [7], [8]. Note that the term "bang–bang" means the single input only switches between two fixed values.

Further assume the bilinear system is time invariant and controllable while the multiplicative terms of the single input and state variables are omitted, i.e., the bilinear system can be treated as a linear time-invariant (LTI) system by neglecting all these multiplicative terms. Based on this LTI system, a novel design technique of sliding function, called the pole-assignment-based method, is presented by directly using the prevailing pole-assignment method [9]. Actually, there are many other techniques for the sliding function design, such as the eigenstructure-assignment method [10] and the Lyapunov-based method [11]. For an LTI system, the Lyapunov-based method is simple and quite straightforward to derive a sliding function, but it is not suitable for a bilinear system due to the fact that the existing multiplicative terms still disturb the system behavior and thus, complicate the process of stability analysis. As for the eigenstructure-assignment method, its main idea is to generate a desired eigenstructure of the sliding mode, just like the pole-assignment-based method introduced in this paper. Therefore, the eigenstructure-assignment method can be also found useful for a bilinear system; however, it is often difficult to achieve an appropriate eigenstructure via the eigenstructure-assignment method. The pole-assignment-based method is then proposed to make it easier to determine the eigenstructure of the sliding mode. In addition, with the help of well-developed tools such as The MATLAB software, the process of the sliding function design is highly simplified in this method.

Since the bang–bang control only can switch between two finite values, the system should be finally restricted to a bounded area. In fact, this area relates to two important regions, called the reaching-and-

Manuscript received March 19, 1998; revised December 1, 1998, September 20, 1999, and May 20, 2000. Recommended by Associate Editor, G. Tao.

Y.-P. Chen and K.-M. Lai are with the Department of Electrical and Control Engineering, National Chiao Tung University, Hsinchu 300, Taiwan.

J.-L. Chang is with the Department of Electrical Engineering, Oriental Institute of Technology, Pan Chiao, Taipei County 220, Taiwan.

Publisher Item Identifier S 0018-9286(00)10016-9.