Proceedings of the 6th WSEAS Int. Conf. on NEURAL NETWORKS, Lisbon, Portugal, June 16-18, 2005 (pp164-169)

Hopfield Networks with Jump Markov Parameters

A.– M. STOICA
Faculty of Aerospace Engineering
University "Politehnica" of Bucharest
Str. Polizu, No. 1
Bucharest, 011061, ROMANIA,

I. YAESH
Control Department
IMI Advanced Systems Div.
P.O.B. 1044/77
Ramat–Hasharon, 47100, ISRAEL

Abstract: – Hopfield networks are symmetric recurrent neural networks which exhibit motions in the state space which converge to minima of energy. Hopfield networks can be used to solve practical complex problems such as implement associative memory, linear programming solvers and optimal guidance problems. In such practical problems, the Hopfield network, may be subject to disturbance signals which can be modelled as finite energy signals. In this paper, we adopt the Lur'e - Postnikov systems approach to analyze Hopfield networks and suggest a training algorithm leading to minimum L_2 gain from the disturbance signals to the error of the network with respect and to its equilibrium points. The suggested algorithm is applied to a numerical example from the field of magnetic heading determination.

Key–Words: – Stochastic H_∞ control, Hopfield Neural Networks, Recurrent Neural Networks, $\mathcal{S}$–procedure, Linear matrix inequalities.

1 Introduction

Hopfield networks are symmetric recurrent neural networks which exhibit motions in the state space which converge to minima of energy. Hopfield networks can be used to solve practical complex problems such as implement associative memory, linear programming solvers and optimal guidance problems. In such practical problems, the Hopfield network, may be subject to exogenous noise which can be modelled as an energy bounded signal. In this paper, we adopt a generalized version of the Lur'e - Postnikov type Lyapunov function (see [1] and [2]) to analyze the L_2 gain of generalized Hopfield networks with Markov jump parameters. The resulting LMI conditions are then applied to analyse the H_∞ norm of the given network or to modify the network weights using a full state feedback controller to achieve minimum noise amplification. This scheme is then used to solve a magnetic heading determination problem from the measured outputs of three axis magnetometer and accelerometers.

The paper is organized as follows. In Section 2, the problem is formulated and in Section 3 Linear Matrix Inequality (LMI) based conditions are derived for L_2 gain analysis of the Hopfield network. These conditions are the basis for an analysis stage where the state space matrices are given. In Section 4, additional LMIs are introduces for controlled networks. Finally Section 5 includes concluding remarks.

Throughout the paper the superscript 'T' stands for matrix transposition, $\mathcal{R}^n$ denotes the n dimensional Euclidean space, $\mathcal{R}^{n\times m}$ is the set of all $n \times m$ real matrices, and the notation $P > 0$, (respectively, $P \geq 0$) for $P \in \mathcal{R}^{n\times n}$ means that P is symmetric and positive definite (respectively, semi-definite). Throughout the paper $(\Omega, \mathcal{F}, \mathcal{P})$ is a given probability space; the argument $\theta \in \Omega$ will be suppressed. Expectation is denoted by $E\{.\}$ and conditional expectation of x on the event $\theta(t) = i$ is denoted by $E[x|\theta(t) = i]$.

2 Problem Formulation

The neural network proposed by Hopfield, can be described by an ordinary differential equation of the form

$$\dot{v}_i(t) = a_i v_i(t) + \sum_{j=1}^{n} b_{ij} g_j(v_j(t)) + c_i = \kappa_i(v), 1 \leq i \leq n \quad (1)$$

where v_i represents the voltage on the input of the ith neuron and where $a_i < 0, 1 \leq i \leq n$ and $b_{ij} = b_{ji}$.

This network in the absence of noise (*i.e.* $w(t) = 0$) is usually analyzed by defining the network energy functional:

$$\begin{aligned} E(v) &= -\sum_{i=1}^{n} a_i \int_0^{v_i} u \frac{dg_i(u)}{du} du \quad (2) \\ &\quad -\frac{1}{2}\sum_{i,j=1}^{n} b_{ij} g_i(v_i) g_j(v_j) - \sum_{i=1}^{n} c_i g_i(v_i) \end{aligned}$$

where it can be seen that $\frac{dE}{dt} = -\sum \frac{dg_i(v_i)}{dv_i} \kappa_i(v)^2 \leq 0$ where the zero rate of the energy is obtained only in the equilibrium points, also referred to as attractors, where

$$\kappa_i(v^0) = 0, 1 \leq i \leq n \quad (3)$$

However, the neural network may be subjected to environmental noise and to connection matrix perturbations. The network subjected to the combination of these two effects can be then described in matrix form as :

$$\dot{v}(t) = Av(t) + Bg(v) + Dw(t) + C, 1 \leq i \leq n \quad (4)$$

where $A := diag(a_1, ..., a_n)$, $B := [b_{ij}]_{i,j=1,...,n}$, $C := \begin{bmatrix} c_1 & c_2 & ... & c_n \end{bmatrix}^T$, $v := \begin{bmatrix} v_1 & v_2 & ... & v_n \end{bmatrix}^T$ and where $g(v) := \begin{bmatrix} g_1(v_1) & g_2(v_2) & ... & g_n(v_n) \end{bmatrix}^T$

The stochastic version of this network driven by white noise, has been considered in [3] the stochastic stability of (1) has been analyzed where it has been shown that the network is almost surely stable when the condition $\frac{dE}{dt} \leq 0$ is replaced by $\mathcal{L}E \leq 0$ where $\mathcal{L}$ is the infinitesimal generator associated with the Ito type stochastic differential equation (4). This condition has been shown in [3] to be satisfied only in cases where the driving noise in (1) is not persistent. This non persistent white noise can be interpreted as a white-noise type uncertainty in A and B but it does not infer any stability results for the practical case of real uncertainties. In the present paper, the Lur'e - Postnikov systems approach ([1],[2]) is invoked to analyze the stability and disturbance attenuation (in the H_∞ norm sense) properties of Hopfield networks subject to Markov jumps in the parameters. The results obtained in this paper reduce to the results in [2] regarding Lur'e - Postnikov systems when the number of jump states is one. The results are given in terms of Linear Matrix Inequalities (LMI) and are applied to a practical example of magnetic heading determination. To analyze the effect of $w(t)$ we first define the error of the Hopfield network output with respect to its equilibrium points by

$$x(t) = v(t) - v^0. \quad (5)$$

The errors vector $x(t)$ satisfy then

$$\dot{x}(t) = Ax(t) + Bf(x) + Dw(t), 1 \leq i \leq n \quad (6)$$

where we assume zero initial conditions, namely

$$x(0) = 0 \quad (7)$$

and where the components $f_k, k = 1, ..., n$ of $f(x) = g(v^0 + x) - g(v^0)$ are assumed to satisfy the sector conditions

$$0 \leq x_k f_k(x_k) \leq x_k^2 \sigma_k \quad (8)$$

which are equivalent to

$$-F_k(x_k, f_k) := f_k(x_k)(f_k(x_k) - \sigma_k x_k) \leq 0 \quad (9)$$

Define the output signals vector of the network to be

$$z(t) = Lx(t). \quad (10)$$

The matrices A, B, C, D and L are piecewise constant matrices of appropriate dimensions whose entries are dependent upon the mode $\theta(t) \in \{1, ..., r\}$ where r is a positive integer denoting the number of possible models between which the Hopfield network parameters can jump. Namely, $A(\theta(t))$ attains the values of $A_1, A_2, ..., A_r$, etc. It is assumed that $\theta(t), t \geq 0$ is a right continuous homogeneous Markov chain on $\mathcal{D} = \{1, ..., r\}$ with a probability transition matrix

$$P(t) = e^{Qt};\ Q = [q_{ij}];\ q_{ii} < 0; \textstyle\sum_{j=1}^{r} q_{ij} = 0; \\ i = 1, 2, .., r\,. \quad (11)$$

Given the initial condition $\theta(0) = i$, at each time instant t, the mode may maintain its current state or jump to another mode $i \neq j$. The transitions between the r possible states, $i \in \mathcal{D}$, may be the result of random fluctuations of the actual network components (*i.e.* resistors, capacitors) characteristics or can used to artificially model deliberate jumps which are the result of parameter changes in an optimization problem the network is used to solve. In the latter case, the transition probabilities, derived from the transition rate matrix Q entries can be used as tuning parameters.

In the next section, we consider the following L_2 gain analysis problem: given $A(\theta(t)), B(\theta(t)), C(\theta(t)), D(\theta(t))$ and $f(x)$ satisfying (8) verify whether the L_2 gain of (6) and (10) is less than $\gamma > 0$, namely the following inequality holds:

$$J = E\left\{\int_0^\infty (z^T(t)z(t) - \gamma^2 w^T(t)w(t))\right\} dt < 0\,. \quad (12)$$

3 L_2 Gain Analysis

The L_2 gain analysis will be based on the Lur'e - Postnikov type ([2]) mode dependent Lyapunov function of the following form :

$$V(x, \theta(t)) = x^T P(\theta(t))x + 2\sum_{k=1}^{n} \lambda_k \int_0^{x_k} f_k(s)ds \quad (13)$$

where $P(\theta(t))$ is a positive definite matrix with r possible values depending on $\theta(t)$and $\lambda_k \geq 0$. The non-quadratic terms multiplying λ_k are aimed at reducing the overdesign involved in the analysis, when $n > 1$.

In the sequel we closely follow a stochastic version of the derivation outlined in [2].

We first note that the *infinitesimal generator* $\mathcal{L}$ associated with $V(x, \theta)$ is then given (see e.g. [4]) by

$$\mathcal{L}V(x, \theta) = QV \\ \quad +diag\left\{ \begin{matrix} h_1^T(x, w)\frac{\partial V}{\partial x} & ... & h_r^T(x, w)\frac{\partial V}{\partial x} \end{matrix} \right\}$$

where $V(x,\theta) := \begin{bmatrix} V(x,1) & \dots & V(x,r) \end{bmatrix}^T$, $\mathcal{L}V(x,\theta) := \begin{bmatrix} \mathcal{L}V(x,1) & \dots & \mathcal{L}V(x,r) \end{bmatrix}^T$ and where $h_i^T(x,w) = x^T A_i^T + f(x)^T B_i^T + w^T D_i^T$.

We define $S = diag\{ \sigma_1 \quad \sigma_2 \quad \dots \quad \sigma_n \}$ where σ_i are the nonlinearity gains of (9). Defining also

$$\Lambda = diag\{ \lambda_1 \quad \lambda_2 \quad \dots \quad \lambda_n \}$$

and

$$T = diag\{ \tau_1 \quad \tau_2 \quad \dots \quad \tau_n \}$$

one can state the main result of this section:

Theorem 1. *The system (6) is stable (in probability) for all x satisfying (9) so that (12) is satisfied if there exist* $0 < P_i \in \mathcal{R}^{n\times n}, i = 1,2,...,r$ *and diagonal matrices* $0 < \Lambda \in \mathcal{R}^{n\times n}$ *and* $0 < T \in \mathcal{R}^{n\times n}$ *that satisfy the following LMI:*

$$\begin{bmatrix} \mathcal{Z}_{i11} & \mathcal{Z}_{i12} & P_i D_i \\ \mathcal{Z}_{i12}^T & \mathcal{Z}_{i22} & \Lambda D_i \\ D_i^T P & D_i^T \Lambda & -\gamma^2 I \end{bmatrix} \le 0, i = 1,2,...,r. \quad (14)$$

where

$$\begin{aligned} \mathcal{Z}_{i11} &:= P_i A_i + A_i^T P_i + \sum_{j=1}^{r} q_{ij} P_j + L_i^T L_i \\ \mathcal{Z}_{i12} &:= P_i B_i + ST + A_i^T \Lambda \qquad (15) \\ \mathcal{Z}_{i22} &:= B_i^T \Lambda + \Lambda B_i - 2T \end{aligned}$$

Proof: Applying the infinitesimal generator $\mathcal{L}$ on $V(x,\theta)$ one obtains that ([4])

$$\mathcal{L}V(x,i) = \sum_{j=1}^{r} q_{ij} V(x,j) + h_i^T(x,w) \frac{\partial V(x,i)}{\partial x} \quad (16)$$

Since from the Ito type formula ([5], [6] and [7]) it results that

$$E\{V(x,\theta(t))|\theta(0)\} = E\{V(x(0),\theta(0))|\theta(0)\} + E\left\{\int_0^t \mathcal{L}V(x(s),\theta(s))ds\right\},$$

it follows that for $x(0) = 0$, (12) is satisfied if

$$\mathcal{L}V \le \gamma^2 w^T w - z^T z. \quad (17)$$

Using the expressions (13) and (16) one obtains that (17) is equivalent with:

$$\begin{aligned} &\left(x^T A_i^T + f^T B_i^T + w^T D_i^T\right)(P_i x + \Lambda f) \\ &+ (x^T P_i + f^T \Lambda)(A_i x + B_i f + D_i w) \\ &+ \textstyle\sum_{j=1}^r q_{ij} x^T P_j x + x^T L_i^T L_i x - w^T w \le 0 \end{aligned}$$

where we have absorbed γ^{-1} in D for simplicity of notations and where (11) and the fact that λ_k do not depend on $\theta(t)$ nulled the extra term $\sum_{j=1}^r q_{ij} \sum_{k=1}^n \lambda_k \int_0^{x_k} f_k(s)ds$. Note that this term would be nonzero if λ_k depended on $\theta(t)$. Completing to squares we get that the inequality $F_{i0}(x,f) \ge 0$ has to be satisfied subject to (9), where:

$$\begin{aligned} &F_{i0}(x,f) := -x^T \left(A_i^T P_i + P_i A_i + L_i^T L_i + P_i D_i D_i^T P_i \right. \\ &\left. + \textstyle\sum_{j=1}^r q_{ij} P_j\right) x - f^T \left(B_i^T \Lambda + \Lambda B + \Lambda D_i D_i^T \Lambda\right) f \\ &- f^T \left(B_i^T P_i + \Lambda A_i + \Lambda D_i D_i^T P\right) x \\ &- x^T \left(P_i B_i + A_i^T \Lambda + P_i D_i D_i^T \Lambda\right) f \\ &+ \left(w^T - x^T P_i D_i - f^T \Lambda D_i\right)\left(w - D_i^T P_i x - D_i^T \Lambda f\right) \end{aligned}$$

Using the $\mathcal{S}$-procedure ([2]) one, therefore, obtains that (17) is satisfied if there exist $\tau_i \ge 0, i = 1,2,...,n$ so that $F_{i0}(x,f) - \sum_{k=1}^n \tau_k F_k(x,f) \ge 0$.

Noticing that

$$\begin{aligned} -\sum_{k=1}^{n} \tau_k F_k(x,f) &= \sum_{k=1}^{n} \left(\tau_k f_k^2 - \tau_k \sigma_k f_k x_k\right) \\ &= f^T T f - f^T \frac{T}{2} S x - x^T \frac{T}{2} S f \end{aligned}$$

it results that (17) is satisfied if

$$\begin{aligned} &x^T \left(A_i^T P_i + P_i A_i + L_i^T L_i \right. \\ &\left. + P_i D_i D_i^T P_i + \textstyle\sum_{j=1}^r q_{ij} P_j\right) x \\ &+ f^T \left(B_i^T \Lambda + \Lambda B_i + \Lambda D_i D_i^T \Lambda - T\right) f \\ &+ f^T \left(B_i^T P_i + S\tfrac{T}{2} + A_i^T \Lambda + \Lambda D_i D_i^T P_i\right) x \\ &+ x^T \left(P_i B_i + S\tfrac{T}{2} + A_i^T \Lambda + P_i D_i D_i^T \Lambda\right) f \le 0 \end{aligned}$$

Substituting T for $\frac{T}{2}$ and applying on the latter Schur complements, the result of the Theorem readily follows reintroducing γ.

Using the above theorem one may establish stability in cases where the system parameters A_i, B_i , etc. jump, due to changes in the optimization objectives of the network, when the Hopfield network continuously evolves its states, without resetting the network.

4 Controlled Hopfield Networks

In many cases, the system parameters A_i, B_i, etc. are given and the corresponding attractor points are the solution of some corresponding optimization problem where (2) is minimized. When the disturbance attenuation γ is not small enough, one may be interested in a subtle modification of the system parameters so as to reduce γ. To this end we introduce a control signal $u := \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix}^T$ into the network in the following manner:

$$\begin{aligned} \dot{x}(t) &= A(\theta(t))x(t) + B(\theta(t))f(x(t)) \\ &\quad + D(\theta(t))w(t) + u(t), \quad u \in \mathcal{R}^n \quad (18) \\ z(t) &= L(\theta(t))x(t) + G(\theta(t))u(t) \end{aligned}$$

where $L_i^T G_i = 0,\ i = 1,...,r$. Find the state feedback control $u(t) = K(\theta(t))x(t)$ such that the resulting system:

$$\begin{aligned} &\dot{x}(t) = (A(\theta(t)) + K(\theta(t)))x(t) \\ &\qquad + B(\theta(t))f(x(t)) + D(\theta(t))w(t) \qquad (19) \\ &z(t) = (L(\theta(t)) + G(\theta(t))K(\theta(t)))x(t) \end{aligned}$$

satisfies the condition in the statement of Theorem 1.

The following result gives necessary and sufficient conditions for the solvability of this problem.

Theorem 2. *The state feedback problem formulated above has a solution if and only if there exist symmetric matrices* $P_i > 0,\ i = 1, ..., r$ *and the diagonal matrices* $\Lambda > 0,\ T > 0$ *verifying the following system of matrix inequalities:*

$$\begin{bmatrix} \mathcal{U}_{11}(P_i, \Lambda, T, S) & 0 & \mathcal{U}_{13}(P_i, \Lambda, T, S) \\ 0 & -\gamma^2 I & -D_i^T G_i^T \\ \mathcal{U}_{13}^T(P_i, \Lambda, T, S) & -G_i D_i & \mathcal{U}_{33}(P_i) \end{bmatrix} < 0 \tag{20}$$

where

$$\begin{aligned} &\mathcal{U}_{11}(P_i, \Lambda, T, S) := P_i^{-1}\left(\textstyle\sum_{j=1}^r q_{ij} P_j\right) P_i^{-1} \\ &\quad + P_i^{-1} L_i^T L_i P_i^{-1} - \Lambda^{-1} T S P_i^{-1} - P_i S T \Lambda^{-1} - 2T \\ &\mathcal{U}_{13}(P_i, \Lambda, T, S) := -P_i^{-1}\left(A_i^T P_i + \textstyle\sum_{j=1}^r q_{ij} P_j\right. \\ &\quad \left. + L_i^T L_i\right) P_i^{-1} G_i^T + \Lambda^{-1} B_i^T G_i^T + \Lambda^{-1} T S P_i^{-1} G_i^T \\ &\mathcal{U}_{33}(P_i) = G_i\left(P_i^{-1} A_i^T + A_i P_i^{-1}\right. \\ &\quad \left. + P_i^{-1}\left(\textstyle\sum_{j=1}^r q_{ij} P_j\right) P_i^{-1} + P_i^{-1} L_i^T L_i P_i^{-1}\right) G_i^T - I \end{aligned}$$

and

$$\begin{bmatrix} B_i^T \Lambda + \Lambda B_i - 2T & \Lambda D_i \\ D_i^T \Lambda & -\gamma^2 I \end{bmatrix} < 0,\ i = 1, ..., r. \tag{21}$$

Proof: Taking into account the orthogonality conditions $L_i^T G_i = 0,\ i = 1, ..., r$ it follows that the conditions of Theorem 1 are fulfilled by the system (19) if there exist the symmetric matrices $P_i > 0,\ i = 1, .., r$ and the diagonal matrices $\Lambda > 0$ and $T > 0$ such that the system of matrix inequalities:

$$\begin{bmatrix} \tilde{\mathcal{Z}}_{i11} + K_i^T G_i^T G_i K_i & \tilde{\mathcal{Z}}_{i12} & P_i D_i \\ \tilde{\mathcal{Z}}_{i12}^T & \tilde{\mathcal{Z}}_{i22} & \Lambda D_i \\ D_i^T P & D_i^T \Lambda & -\gamma^2 I \end{bmatrix} \le 0, \quad i = 1, 2, ..., r \tag{22}$$

where

$$\begin{aligned} \tilde{\mathcal{Z}}_{i11} &:= P_i(A_i + K_i) + (A_i + K_i)^T P_i \\ &\quad + \sum_{j=1}^r q_{ij} P_j + L_i^T L_i \\ \tilde{\mathcal{Z}}_{i12} &:= P_i B_i + ST + (A_i + K_i)^T \Lambda \\ \tilde{\mathcal{Z}}_{i22} &:= B_i^T \Lambda + \Lambda B_i - 2T \end{aligned} \tag{23}$$

$i = 1, ..., r$ is feasible. Based on a Schur complement argument the inequalities (22) may be rewritten in the equivalent form:

$$\begin{bmatrix} \tilde{\mathcal{Z}}_{i11} & \tilde{\mathcal{Z}}_{i12} & P_i D_i & K_i^T G_i^T \\ \tilde{\mathcal{Z}}_{i12}^T & \tilde{\mathcal{Z}}_{i22} & \Lambda D_i & 0 \\ D_i^T P & D_i^T \Lambda & -\gamma^2 I & 0 \\ G_i K_i & 0 & 0 & -I \end{bmatrix} \le 0, \quad i = 1, 2, ..., r. \tag{24}$$

In the following the case when the above inequalities are strict will be considered. Further the inequalities (24) can be written as:

$$\mathcal{Z}_i + \mathcal{P}_i^T K_i \mathcal{Q}_i + \mathcal{Q}_i^T K_i^T \mathcal{P}_i < 0,\ i = 1, ..., r \tag{25}$$

where:

$$\begin{aligned} \mathcal{Z}_i &:= \begin{bmatrix} \mathcal{Z}_{i11} & \mathcal{Z}_{i12} & P_i D_i & 0 \\ \mathcal{Z}_{i12}^T & \mathcal{Z}_{i22} & \Lambda D_i & 0 \\ D_i^T P_i & D_i^T \Lambda & -\gamma^2 I & 0 \\ 0 & 0 & 0 & -I \end{bmatrix} \\ \mathcal{P}_i &:= \begin{bmatrix} P_i & \Lambda & 0 & G_i^T \end{bmatrix} \\ \mathcal{Q}_i &:= \begin{bmatrix} I & 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

and $\mathcal{Z}_{i11}$, $\mathcal{Z}_{i12}$ and $\mathcal{Z}_{i22}$ are defined in (15). Then the Projection Lemma (see *e.g.* [2]) gives that (25) are feasible if and only if:

$$W_{\mathcal{P}_i}^T \mathcal{Z}_i W_{\mathcal{P}_i} < 0 \tag{26}$$

and

$$W_{\mathcal{Q}_i}^T Z_i W_{\mathcal{Q}_i} < 0, \tag{27}$$

where $W_{\mathcal{P}_i}$ and $W_{\mathcal{Q}_i}$are bases of the null spaces of $\mathcal{P}_i$ and $\mathcal{Q}_i$, respectively. Then one can directly see that

$$W_{\mathcal{P}_i} = \begin{bmatrix} P_i^{-1}\Lambda & 0 & -P_i^{-1} G_i^T \\ -I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}$$

and thus condition (20) directly follows after pre and post-multiplication of (26) by *diag* $\left(\Lambda^{-1},\ I,\ I\right)$. Further, since

$$W_{\mathcal{Q}_i} = \begin{bmatrix} 0 & 0 & 0 \\ I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix},$$

direct computations give that condition (27) is equivalent with (21)and thus the proof ends.

Remark 1. *If the systems of matrix inequalities (20) and (21) are feasible then the state feedback gains* K_i *are obtained solving the linear matrix inequalities (25).*

The inequalities (20) cannot be written in an equivalent linear form due to the terms $\Lambda^{-1} T S P_i^{-1}$ appearing in the elements (1,1) and (1,3). The idea to introduce an additional variable ([2]) $Z_i = \Lambda^{-1} T S P_i^{-1}$ is not effective since $\Lambda^{-1} T S$ must result a diagonal matrix. Then an alternative numerical method must be used.

Assuming that the conditions of the above theorem are fulfilled, the gains $K_i,\ i = 1, ..., r$ can be determined as follows:

Step 1. Solve the system of LMIs (21) obtaining thus the diagonal matrices $\Lambda > 0$ and $T > 0$; one can see that the system (21) is also checked by any other diagonal matrix larger than T.

Proceedings of the 6th WSEAS Int. Conf. on NEURAL NETWORKS, Lisbon, Portugal, June 16-18, 2005 (pp164-169)

Step 2. Taking $P_i = \frac{1}{\varepsilon}I$ and $T = \frac{1}{\varepsilon}I$ with $\varepsilon > 0$ the condition (20) becomes:

$$\begin{bmatrix} \mathcal{V}_{11}(\varepsilon) & 0 & \mathcal{V}_{13}(\varepsilon) \\ 0 & -\gamma^2 I & -D_i^T G_i^T \\ \mathcal{V}_{13}^T(\varepsilon) & -G_i D_i & \mathcal{V}_{33}^T(\varepsilon) \end{bmatrix} < 0, \;\; i = 1, ..., r \tag{28}$$

where

$$\begin{aligned} \mathcal{V}_{11}(\varepsilon) &:= \varepsilon^2 L_i^T L_i - \Lambda^{-1} S - S\Lambda^{-1} - \frac{2}{\varepsilon} \\ \mathcal{V}_{13}(\varepsilon) &:= -\varepsilon\left(A_i^T + \varepsilon L_i^T L_i\right) G_i^T + \Lambda^{-1} B_i^T G_i^T \\ &\quad +\Lambda^{-1} S G_i^T \\ \mathcal{V}_{33}(\varepsilon) &:= \varepsilon G_i\left(A_i^T + A_i + \varepsilon L_i^T L_i\right) G_i^T - I \end{aligned}$$

Inspecting (28) it follows that these inequalities are fulfilled for $\varepsilon > 0$ small enough if:

$$\gamma > \max_{i \in \{1,...,r\}} \rho^{\frac{1}{2}}\left(D_i^T G_i^T G_i D_i\right) \tag{29}$$

with $\rho(.)$ denoting the spectral radius. Thus with $P_i = T = \frac{1}{\varepsilon}I$ determined for a small enough $\varepsilon > 0$ and with Λ obtained at Step 1, one can solve the basic LMIs (25) with respect to $K_i, \;\; i = 1, ..., r$.

5 Example

We consider the problem of obtaining the azimuth angle from 3-axis magnetometer, given its readings of the components of the magnetic field of Earth and noisy measurements the pitch and roll angles. It is assumed that the azimuth angle is constant, for example a camera mounted on a car travelling on a straight road for some time. Denoting the Euler angles of the camera axis by ψ, θ, ϕ (namely 3 successive rotations transform Earth fixed axes to camera fixed axes), the noisy measurements of roll and pitch angles can be, for example, obtained from accelerometer readings where they noise can be the results of road imperfectness such as road bumps. Neglecting this maneuver noise (which will be re-introduced to the simulations), the accelerometers measure

$$\begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} = \begin{bmatrix} g sin(\theta) \\ -g sin(\phi) cos(\theta) \\ -g cos(\phi) cos(\theta) \end{bmatrix}$$

where g is the local gravity constant. The angles θ and ϕ are then derived from $\theta = sin^{-1}(a_x/g)$ and $\phi = -sin^{-1}(a_y/cos(\theta))$. We denote the magnetic field vector in the North-East-Down frame by $e := \begin{bmatrix} e_1 & e_2 & e_3 \end{bmatrix}^T$ and the magnetometer readings $m := \begin{bmatrix} m_1 & m_2 & m_3 \end{bmatrix}^T$. These two vectors are related by $m = \mathcal{L}v$ where $\mathcal{L}(\psi, \theta, \phi)$ is the direction cosines matrix (DCM) corresponding to ψ, θ, ϕ (see eq. (4.54) in [10]). Since we define the azimuth to be relative to the magnetic north of Earth, we just take $e := \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$. Denoting also $x := \begin{bmatrix} cos(\psi) & sin(\psi) \end{bmatrix}^T$ and

$$b^{(i)} := \begin{bmatrix} m_1 + sin(\theta^{(i)})e_3 \\ m_2 - sin(\phi^{(i)})cos(\theta^{(i)})e_3 \\ m_3 - cos(\phi^{(i)})cos(\theta^{(i)})e_3 \end{bmatrix}$$

we readily obtain that each three axis magnetometer reading $m^{(i)}$ is related to $cos(\psi)$ and $sin(\psi)$ solution to the problem of resolving the azimuth angle ψ out of the magnetometer measurements corresponds to $b^{(i)} = A^{(i)}x$ where

$$A^{(i)} := \begin{bmatrix} cos(\theta^{(i)})e_1 & cos(\theta^{(i)})e_2 \\ s_{\phi,\theta}(i)e_1 + cos(\phi^{(i)})e_2 & -cos(\phi^{(i)})e_1 + s_{\phi,\theta}(i)e_2 \\ c_{\phi,\theta}(i)e_1 - sin(\phi^{(i)})e_2 & sin(\phi^{(i)})e_1 + c_{\phi,\theta}(i)e_2 \end{bmatrix}$$

with $s_{\phi,\theta}(i) := sin(\theta^{(i)})sin(\phi^{(i)})$, $c_{\phi,\theta}(i) := cos(\phi^{(i)})sin(\theta^{(i)})$ and where x should satisfy the constraint $x^T x = 1$. Since the camera is assumed to be stationary $i = 1, 2...N$ measurements can be taken. In such a case, defining $b = \begin{bmatrix} b^{(1)} \\ b^{(2)} \\ ... \\ b^{(N)} \end{bmatrix}$ and $A = \begin{bmatrix} A^{(1)} \\ A^{(2)} \\ ... \\ A^{(N)} \end{bmatrix}$ the problem of resolving ψ from the N measurements becomes one of solving the following normalized least-squares problem to minimize $J = (b^T - x^T A^T)(b - Ax)$ subject to the constraint $x^T x = 1$. Note that $A \in \mathcal{R}^{N,2}$ where $N >> 1$ and without the constraint $x^T x = 1$ a solution to this problem would be $x = (A^T A)^{-1} A^T b$. To impose the constraint $x^T x = 1$ on the solution we suggest an iterative solution which makes use of the Hopfield network concept with Markov jump parameters. To this end, first note that J can be written as

$$J_1 = x^T A^T A x - 2x^T A^T b + b^T b \tag{30}$$

Defining $x_i = g_i(v_i)$ where $g(.) = tan^{-1}(.)$ and defining $b_{i,j} = -2\{A^T A\}_{i,j} \quad c_i = 2\{A^T b\}_i$ we notice that (30) can be written as (2) for small $|a_i|$. To impose the constraint $x^T x = 1$ one would need to add to J of (30) the term $(1 - x^T x)^2$ which is biquadratic in x rather than quadratic to fit the form of (2). To resolve this problem we may use the observation that as x will approach it steady state solution (*i.e.* the local minimum of the energy function of (2), one may be able to approximate $(1 - x^T x)^2$ by $(1 - x^T x_{prev})^2$ where x_{prev} is the previous value of x during the recursions to numerically solve (1) say by Euler integration. We, therefore define also $J_2 = x^T A^T A x - 2x^T A^T b + b^T b + \rho(1 - x^T x_{prev})^2$. We see that $J_2 = x^T A^T A x - 2x^T A^T b + b^T b + \rho(1 + x^T x_{prev} x_{prev}^T x - 2x^T x_p rev)$. Namely,

$J_2 = x^T(A^TA+\rho x_{prev}x_{prev}^T)x-2x^T(A^Tb+\rho x_{prev})+b^Tb+\rho$. Comparing the latter with (2) we choose $B_1 = -2A^TA$, $B_2 = -2A^TA - 2\rho x_{prev}x_{prev}^T$, $C_1 = -2A^Tb$ and $C_2 = 2A^Tb + 2\rho x_{prev}$. We also choose $A = -\bar{\epsilon}I$ for small $\bar{\epsilon} > 0$ so that it has a negligible effect on (2). The solution to our problem will be, therefore, obtained by applying (4) on the above chosen parameters where we choose $\bar{\epsilon} = 0.01$ and where $Q = 0.05\begin{bmatrix} -0.5 & 0.5 \\ 0.5 & -0.5 \end{bmatrix}$ defines the transitions rate between $r = 1$ and $r = 2$ affecting A and B. Since $\mathcal{P}(t) = e^{Qt}$ and where the corresponding infinitesimal matrix is $Q = \begin{bmatrix} -\alpha & \alpha \\ \beta & -\beta \end{bmatrix}$ where $\alpha = \beta = 0.025$. It follows that $\mathcal{P}(t) = e^{Qt} = (\alpha t+\beta t)^{-1}\begin{bmatrix} \beta t+\mu(t)\alpha t & \alpha t-\mu(t)\alpha t \\ \beta t-\mu(t)\beta t & \alpha t+\mu(t)\beta t \end{bmatrix}$ where $\mu(t) = e^{-(\alpha+\beta)t}$. Namely, near $t = 0$ the mode $\eta(t)$ will spend more time at $\eta(t) = 1$ and as t increases more time will be given to $\eta(t) = 2$ for the sake of satisfying $x^Tx = 1$.

We next simulate our algorithm where we take true values of the roll and pitch angles to be $\phi = -15$ degrees and $\theta = 10$ degrees. We also take the true magnetic azimuth to be $\psi = 145$ degrees. The roll and pitch angles are measured with additive Gaussian noise of zero mean and 10 degrees standard deviation, and the magnetometer output is taken without noise. Choosing $\rho = 10$ we obtain the simulation results,obtained with an integration step of $\Delta t = 0.01$ second and Euler method for integration. The convergence of $cos(\psi)$ and $sin(\psi)$ has been obtained, and the fact that the constraint tended to convergence during periods where $i = 2$ was observed. The solution for ψ_0 the Hopfield network converged to the minimum of the cost function J. The fact that the stochastic Hopfield system we simulated is stochastically stable has been verified by solving (22) of Theorem 1 using [9]. We obtained $P_1 = \begin{bmatrix} 0.0309 & 0.0000 \\ 0.0000 & 0.0310 \end{bmatrix}$, $P_2 = \begin{bmatrix} 0.0255 & 0.0036 \\ 0.0036 & 0.0283 \end{bmatrix}$, $T = 10^{-3}\begin{bmatrix} 1.1723 & 0 \\ 0 & 0.6134 \end{bmatrix}$ and $\Lambda = \begin{bmatrix} 10^{-3}\cdot 0.7642 & 0 \\ 0 & 0.5995 \end{bmatrix}$ all being positive definite matrices and disturbance attenuation factor $\gamma = 0.0313$.

6 Conclusions

A class of stochastic Hopfield networks where the network weights jump according a Markov chain process have been considered. Both stochastic stability and disturbance attenuation analysis in an H_∞ setup have been related to Linear Matrix Inequalities which are easy to solve. A class of controlled stochastic Hopfield networks has been introduced where the disturbance attenuation factor can be modified via state-feedback. The theory of stochastic stability and disturbance attenuation analysis has been applied to a simple problem from the field of magnetic heading determination. Although the problem is simple, it may represent more involved problems where quaternions have to be estimated from some related measurements ([8]) where normalization constraints should be applied to the estimates. The problems of reducing the inherent overdesign in the synthesis of state-feedback modification of the Hopfield networks and exploring its possible applications is left for future research.

References:

[1] A.I. Lure and V.N. Postnikov " On the theory of stability of control systems", *Applied Mathematics and Mechanics*, Vol. 8, No. 3, 1944, in Russian.

[2] S. Boyd, El. Ghaoui, L. Feron and V. Balakrishnan, *Linear matrix inequalities in system and control theory* Philadelphia, PA:SIAM, 1994.

[3] S. Hu, X. Liao and X. Mao, "Stochastic Hopfield Neural Networks", Statistical pulse compression, *Journal of Physics A: Mathematical and General*, Vol. 36, pp. 1-15, 2003.

[4] X. Fen, K.A. Loparo, Y. Ji and H.J. Chizeck. "Stochastic Stability Properties of Jump Linear Systems", *IEEE Transactions on Automatic Control*, Vol. 37, pp. 38-53,1992.

[5] V. Dragan and T. Morozan, "Stability and Robust Stabilization to Linear Stochastic Systems Described by Differential Equations with Markovian Jumping and Multiplicative White Noise", Preprint 17/1999, The Institute of Mathematics of the Romanian Academy, 1999.

[6] V. Dragan and T. Morozan, "Stability and robust stabilization to linear stochastic systems described by differential equations with Markovian jumping and multiplicative noise", *Stoch. Anal. Applicat.*, Vol. 22, No. 1, pp/ 22-92,2002.

[7] V. Dragan and T. Morozan, "The linear quadratic optimiziation problems for a class of linear stochastic systems with multiplicative white noise and Markovian jumping", *IEEE Transactions on Automat. Contr.*, Vol. 49, No. 5, pp. 665 - 675, 2004.

[8] J. L. Marins, X. Yun, E. R. Bachmann, R. McGhee and M.J. Zyda, "An Extended Kalman FIlter for Quaternion-Based Orientation Estimation Using MARG Sensors", Proceedings of the 2001 IEEE/RSJ, International Conference on Intelligent Robots and Systems, Maui, Hawaii, USA, October 29-November 03, 2001.

[9] P. Gahinet, A. Nemirovsky, A. J. Laub and M. Chilali, "LMI Control Toolbox for Use with MATLAB ", The Mathworks Inc. 1995.

[10] B. Etkin, *Dynamics of Athmospheric Flight*, John Wiley & Sons, New-York, 1972.