2008 American Control Conference
Westin Seattle Hotel, Seattle, Washington, USA
June 11-13, 2008

Feedback Error Learning with a Noisy Teacher

Abraham K. Ishihara, Johan van Doornik, and Shahar Ben-Menahem

Abstract— **The Feedback Error Learning (FEL) algorithm is examined under the condition of a noisy teaching signal. The teaching signal, which adaptively adjusts the weights of the feedforward network, is assumed to be corrupted by a Signal Dependent Noise (SDN) source. The FEL framework was originally inspired by the cerebellum as a model for human motor control. We analyze the robustness properties of the original system with respect to the SDN noise model. We prove bounds on the learning rate and feedback gain matrices that guarantee stochastic stability of the closed loop system.**

I. INTRODUCTION

The Feedback Error Learning framework was originally proposed by Kawato [1] as a first attempt to link the neuro-anatomical findings in the cerebellum and motor cortex with components of an adaptive control scheme. Kawato claimed that the training signal which originates in the inferior olive in the medula region of brain stem conveys information of the feedback control signal via its climbing fibers. The climbing fibers project onto the dendrites of Purkinje cells and are known to induce Long Term Depression (LTD) which is the primary mechanism that modulates cerebellar output. Later, Garwicz [4], showed by microelectrode stimulation and recording in the inferior olive of a cat that the climbing fiber input to Purkinje cells may encode a motor command derived from a spinal feedback loop. In other words, the teaching signal may simply encode the feedback control signal. The FEL architecture is depicted in Fig. 1. More recent theoretical studies of the FEL framework are given in [2] and [3].

In this note, we will examine the signal dependent noise (SDN) model that has recently received much attention in the neuroscience community [5],[6]. It has long been known that neural signals are highly variable. Whether this variability is the result of an additive noise source or part of the signal is unknown [7]. Harris and Wolpert [8] proposed that variability of neural signals during the control of eye and arm movements is well accounted for by a neural noise model in which the standard deviation of the control signal increases linearly with the mean of the signal. In other words, the additive noise is signal dependent on the signal being corrupted. More importantly, this hypothesis is consistent with empirical observations of the trade off between speed and accuracy [9]. Using the SDN model, Harris and Wolpert were able to predict the bell shaped velocity profiles observed experimentally. It is still unclear where and what signals in the nervous system are being corrupted by signal dependent noise. Potential candidates include higher planning areas in the brain such as premotor or supplementary motor areas, motor command signals descending to the spinal level via the corticospinal tract or rubrospinal tract, or the various sensory feedback pathways, such as the dorsal column medial lemniscous pathway. In this note, we analyze the case where signal dependent noise corrupts the teaching signal of the feedforward network. In the FEL framework, this corresponds to the signal generated from inferior olive. The signal dependent noise corruption is indicated by the block labeled $\sigma\dot{B}$ in Fig. 1. A more detailed description will be given below.

A. Ishihara is with Carnegie-Mellon University West Campus, Moffett Field, CA 94035 ishihara@stanford.edu

J. van Doornik is with the Div. Child Neurology and Movement Disorders Stanford University Medical Center 300 Pasteur, room A345 Stanford, CA 94305-5235 USA jvd@stanford.edu

S. Ben-Menahem is with the Department of Physics, Stanford University, Stanford CA 94305, U.S.A, and Avago Technologies Inc., San Jose CA USA sanverim@gmail.com

II. STOCHASTIC SETTING

In this section we define the relevant stochastic terminology as well as Ito's formula. We refer the reader to [11]. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ be a complete probability space. Let B_t be an m-dimensional Brownian motion vector adapted to the filtration, $\{\mathcal{F}_t\}$. An d-dimensional Ito process, $x(t,\omega)$, is a continuous, $\{\mathcal{F}_t\}$-adapted stochastic process that is the solution of the stochastic differential equation denoted by

$$dx(t,\omega) = f(x(t,\omega),t)dt + g(x(t,\omega),t)dB$$

or in other words, $x(t,\omega)$ satisfies the equality

$$x(t,\omega) = x(0,\omega) + \int_0^t f(x(s,\omega),s)ds + \int_0^t g(x(s,\omega),s)dB$$

for all $t \in [0,T]$. In the following, we will suppress the dependence of x on $\omega \in \Omega$ and denote $x(t,\omega)$ as simply $x(t)$. It is to be understood that for each fixed t, x is a random variable. Let $\mathcal{L}^p(X;Y)$ denote the collection of stochastic processes $h : X \times \Omega \to Y$ such that

$$\int_0^T \|h(s)\|_Y^p ds < \infty \quad a.s \;\; \forall T > 0$$

We now state the Ito formula, which is the stochastic version of the chain rule of ordinary calculus.

Theorem 1 (Ito's Lemma-adapted from [11]): Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ be a complete probability space, and $x(t,\omega)$ be a d-dimensional Ito process satisfying

$$x(t,\omega) - x(0,\omega) = \int_0^t f(x,s)ds + \int_0^t g(x,s)dB \qquad (1)$$

where $f \in \mathcal{L}^1(\mathbb{R}^n \times \mathbb{R}^+ \times \Omega; \mathbb{R}^d)$, $g \in \mathcal{L}^2(\mathbb{R}^n \times \mathbb{R}^+ \times \Omega; \mathbb{R}^{d\times m})$, and $B : \mathbb{R}^+ \times \Omega \to \mathbb{R}^m$ is an m dimensional

978-1-4244-2079-7/08/$25.00 ©2008 AACC.

Brownian motion. Let $V \in \mathrm{C}^{2,1}(\mathbb{R}^d \times \mathbb{R}^+; \mathbb{R})$. Then,

$$V(x(t),t) - V(x(0),0) = \int_0^t \bar{f} ds + \int_0^t \bar{g} dB \quad \text{a.s}$$

where $\bar{f} = V_t + V_x f + \frac{1}{2}\mathrm{Tr}\left(g^T V_{xx} g\right)$ and $\bar{g} = V_x g$.
Note the additional term given by

$$\frac{1}{2}\mathrm{Tr}\left(g^T V_{xx} g\right) \tag{2}$$

This is the Ito *anomaly* and is the reason why stability analysis of stochastic systems is very different than its deterministic counterpart.

Definition 1 (Stochastic Stability): The equilibrium of (1) is said to be *stochastically stable*, or *stable in probability* if for every pair of $\epsilon \in (0,1)$ and $r > 0$, there exists a $\delta = \delta(\epsilon, r, t_0)$ such that $\|x_0\| < \delta$ implies

$$P\{\|x(t;t_0;x_0)\| < r \quad \forall t \geq t_0\} \geq 1 - \epsilon$$

For brevity, we define the operator L acting on the Lyapunov function V as

$$LV := V_t + V_x f + \frac{1}{2}\mathrm{Tr}\left(g^T V_{xx} g\right)$$

We may now succinctly write Ito's Lemma for the function $V \in \mathrm{C}^{2,1}(\mathbb{R}^d \times \mathbb{R}^+; \mathbb{R})$ as

$$dV = LV dt + V_x g(x,t) dB$$

Theorem 2: (*Adapted from [11]*): Let $S \subset \mathbb{R}^d$. If there exists a positive definite function $V \in \mathrm{C}^{2,1}(S \times [0,\infty); \mathbb{R})$ such that $LV \leq 0$ for all $(x,t) \in S \times [0,\infty)$, then the equilibrium of (1) is *stochastically stable*.

III. Background

In this section, we state the plant dynamics, the feedforward control, and the FEL algorithm for the deterministic case originally presented in [1]. We take the plant to be a serial link manipulator. The dynamics are given as follows:
Plant Dynamics

$$M(q)\ddot{q} + V(q,\dot{q}) + F(\dot{q}) + G(q) + \tau_d = \tau \tag{3}$$

where $M(q) \in R^{n \times n}$ is the inertia matrix, $V(q,\dot{q}) \in R^{nx1}$ is the corriolis/centripetal matrix, $G(q) \in R^{nx1}$ is the gravity matrix, $F \in \mathbb{R}^n$ denotes friction terms, and $\tau_d \in R^{nx1}$ represents unknown disturbances. $q \in R^{nx1}$ is the joint angle state vector, and $\tau \in R^{nx1}$ is the net torque applied at each joint. We have the following standard assumptions [14].

Assumption 1: M(q) is symmetric, positive definite and bounded.

Assumption 2: The corolis/centripetal terms can be written asrolis/centripetal terms can be written as $V_m(q,\dot{q})\dot{q}$ such that $\dot{M} - 2V_m(q,\dot{q})$ is skew-symmetric.

Assumption 3: The gravity and disturbance terms are bounded.

We will neglect friction and disturbance terms in the analysis. The desired trajectory is denoted by $\vec{q_d} = \begin{bmatrix} q_d \\ \dot{q}_d \\ \ddot{q}_d \end{bmatrix}$. The control input to the plant is denoted by $\tau = \tau_{fb} + \tau_{ff}$. The feedback control is given by $\tau_{fb} = K_p \tilde{q} + K_v \dot{\tilde{q}}$ where $\tilde{q} := q_d - q$, and K_s for $s = p, v$ denotes a positive definite symmetric feedback gain matrix. We will use the following notation for the lower and upper bounds of key terms:

$$\begin{array}{ll} \kappa_M \stackrel{\mathsf{d}}{=} \inf_{q \in \mathbb{R}^n} \sigma_{min}(M(q)) & \omega_M \stackrel{\mathsf{d}}{=} \sup_{q \in \mathbb{R}^n} \|M(q)\| \\ \kappa_v \stackrel{\mathsf{d}}{=} \sigma_{min}(K_v) & \omega_v \stackrel{\mathsf{d}}{=} \|K_v\| \\ \kappa_p \stackrel{\mathsf{d}}{=} \sigma_{min}(K_p) & \omega_p \stackrel{\mathsf{d}}{=} \|K_p\| \\ \beta_p = \frac{\omega_p}{\kappa_p} & \beta_v = \frac{\omega_v}{\kappa_v} \end{array}$$

The feedforward command, τ_{ff}, is the output of a basis function network, for example a radial basis function (RBF) network:

$$\begin{aligned} \tau_{ff} &= \hat{W}^T \phi(q_d) \\ &= \left[W^* - \tilde{W}\right] \phi(q_d) \\ &= \tau^*_{ff} - \epsilon(q_d) - \tilde{W}^T \phi(q_d) \end{aligned} \tag{4}$$

where $\hat{W} \in \mathbb{R}^{N \times n}$ is a matrix of estimated parameters, and $\tilde{W} \in \mathbb{R}^{N \times n}$ denotes the matrix of parameter errors defined by $\tilde{W} \stackrel{\mathsf{d}}{=} W^* - W$. Note that we have assumed that there is no approximation error in the network. In other words, we have assumed that we have knowledge of the exact basis function structure (the regressor). An exact basis always exists for the n-link, rigid, robot manipulator [14]. The case for which approximation error exists is handled in a forthcoming paper, and requires the addition of Narendra's e-modification to the FEL algorithm.

The FEL algorithm [1] is given as

$$\dot{\hat{W}} = \Gamma \phi(q_d) \tau_{fb}^T \tag{5}$$

where $\Gamma \in \mathbb{R}^{N \times N}$ is a symmetric, positive definite matrix of learning rates. This algorithm can be interpreted as a continuous least mean square algorithm with respect to the inverse dynamics of the plant assuming that the desired output is the actual control input to the plant. We do not go into details here.

IV. Main Results

In this section, we consider signal dependent noise on the signal that trains the feedforward network. In physiological terms, this may apply to climbing fiber excitation of Purkinje cells in lateral cerebellum. With respect to the FEL control system, this noise source is modeled as shown in Fig. 1.

Here, we assume that the training signal to the network is corrupted by signal dependent noise, in which each component of the signal is multiplied by $1 + \sigma_i \dot{B}_i$, where σ_i represents the noise intensity of the ith component of the Brownian motion.

Before proceeding with the system above, we consider a somewhat general SDE and derive an expression for the Ito anomaly. To this end, consider an SDE of the form:

$$\dot{x} = F(x,t) + G(x,t)\dot{B} \tag{6}$$

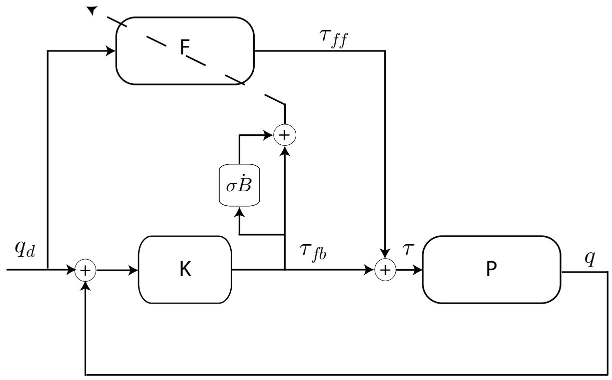

Fig. 1. FEL block diagram with signal dependent noise corrupting the training signal to the feedforward neural network.

where

$$x = \begin{bmatrix} x^{(1)} \\ \vdots \\ x^{(d)} \end{bmatrix} \qquad F(x,t) = \begin{bmatrix} F_1(x,t) \\ \vdots \\ F_d(x,t) \end{bmatrix}$$

$$G(x,t) = \begin{bmatrix} G_1(x,t) \\ \vdots \\ G_d(x,t) \end{bmatrix} \qquad B(t) = \begin{bmatrix} B_1(t) \\ \vdots \\ B_m(t) \end{bmatrix}$$

and the dimensions of the components are $x^{(i)} \in \mathbb{R}^{n_i}$, $F_i \in \mathbb{R}^{n_i}$, $G_i \in \mathbb{R}^{n_i \times m}$ and $\sum n_i = \bar{n}$. As before, the components of B are scalar Brownian motions.

We now consider the Ito anomaly given in (2). Due to symmetry, the Ito anomaly is equivalent to $\frac{1}{2}\text{Tr}\left(V_{xx}GG^T\right)$. It can be shown that

$$\text{Tr}\left(V_{xx}GG^T\right) = \sum_{i,j=1}^{d} \text{Tr}\left(V_{x^{(j)}x^{(i)}} G_i G_j^T\right) \tag{7}$$

where

$$V_{x^{(k)}x^{(j)}} := \begin{bmatrix} \frac{\partial}{\partial x_1^{(k)}} \frac{\partial V}{\partial x_1^{(j)}} & \cdots & \frac{\partial}{\partial x_1^{(k)}} \frac{\partial V}{\partial x_{n_j}^{(j)}} \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_{n_k}^{(k)}} \frac{\partial V}{\partial x_1^{(j)}} & \cdots & \frac{\partial}{\partial x_{n_k}^{(k)}} \frac{\partial V}{\partial x_{n_j}^{(j)}} \end{bmatrix} \in \mathbb{R}^{n_k \times n_j}$$

We now consider the learning dynamics corrupted by SDN on the training signal. Each component of the feedback, τ_{fb_i} is corrupted by a signal dependent brownian motion with noise intensity, σ_i. The resulting stochastic differential equation is given by

$$\dot{\tilde{W}} = -\gamma\phi(q_d)\tau_{fb}^T - \gamma\phi(q_d)(\Lambda\dot{B})^T \tag{8}$$

where $\tilde{W} = W^* - \hat{W}$ denotes the $N \times n$ matrix of weight errors and $\phi(q_d) \in \mathbb{R}^N$ denoted the RBF network with the desired trajectory as the input. The stochastic terms are defined as $\Lambda = \text{diag}\{\sigma_i \tau_{fb_i}\} \in \mathbb{R}^{n \times n}$ and $B = [B_1 B_2 \dots B_m]^T$ denotes an m-dimensional Brownian motion.

Closed Loop System Dynamics: We now vectorize the plant (3) and learning (8) dynamics. Let $x^{(1)} := \tilde{q}$ and $x^{(2)} = \dot{\tilde{q}}$. Vectorizing the weight matrix, we let

$$x^{(3)} = \left[\tilde{W}_{[1,:]} \; \tilde{W}_{[2,:]} \cdots \tilde{W}_{[N,:]}\right]^T \in \mathbb{R}^{N \cdot n}$$

where $\tilde{W}_{[i,:]}$ denotes the ith row of matrix $\tilde{W}$. Hence, the closed loop system is described by

$$\begin{aligned} \dot{x}^{(1)} &= F_1(x,t) + G_1(x,t)\dot{B} \\ \dot{x}^{(2)} &= F_2(x,t) + G_2(x,t)\dot{B} \\ \dot{x}^{(3)} &= F_3(x,t) + G_3(x,t)\dot{B} \end{aligned} \tag{9}$$

where

$$\begin{aligned} F_1(x,t) &= x^{(2)} \\ F_2(x,t) &= \ddot{q}_d - M^{-1}(x^{(1)})\left(\tau - V_m x^{(2)} - G\right) \end{aligned}$$

$$F_3(x,t) = -\Gamma \begin{bmatrix} \phi_1(\mathbf{q_d})\tau_{fb} \\ \phi_2(\mathbf{q_d})\tau_{fb} \\ \vdots \\ \phi_N(\mathbf{q_d})\tau_{fb} \end{bmatrix} \quad G_3(x,t) = -\Gamma \begin{bmatrix} \phi_1(\mathbf{q_d})\Lambda \\ \phi_2(\mathbf{q_d})\Lambda \\ \vdots \\ \phi_N(\mathbf{q_d})\Lambda \end{bmatrix}$$

and $G_1 = G_2 = 0$. Before stating the main theorem, we define a key constant that is used throughout the paper. Define the constant

$$b_0 := \sigma_B^2 \phi_B^2 \tag{10}$$

where σ_B is the maximum component of the noise intensity vector defined by

$$\sigma_B := \max_{i \in [1,n]} \{|\sigma_i|\} \tag{11}$$

and ϕ_B is the supremum norm of the basis function network over all possible desired trajectories defined by

$$\phi_B := \sup_{q_d \in \mathcal{K}} \|\phi(q_d)\| \tag{12}$$

Theorem 3: Consider the closed loop system given in (9). Let the ideal feedforward component be defined by

$$\tau_{ff}^* = M(q_d)\ddot{q}_d + V_m(q_d, \dot{q}_d)\dot{q}_d + G(q_d) = W^{*T}\phi(q_d)$$

for all $\begin{bmatrix} q_d \\ \dot{q}_d \\ \ddot{q}_d \end{bmatrix} \in \mathcal{K}$, where $\mathcal{K}$ is a known compact set. Let the control input to the plant be denoted by $\tau = \tau_{fb} + \tau_{ff} \in \mathbb{R}^n$ where $\tau_{fb} = K_p\tilde{q} + K_v\dot{\tilde{q}}$, and $\tau_{ff} = \hat{W}^T\phi(q_d)$. Let the network be trained according to the FEL algorithm with SDN corrupting the training signal as given in (8). Let c_γ be an arbitrary positive real number. Let the triple $(\kappa_p, \kappa_v, \gamma)$ be selected in accordance with following bounds:

$$\kappa_p > \frac{1}{2}\left(v_3 + \sqrt{v_3^2 + \alpha_2^2}\right)$$

$$\kappa_v > \max\left(\frac{d_2 + \sqrt{d_2^2 + 4d_1d_3}}{2d_1}, \omega_p\omega_M\sqrt{\frac{1}{2\kappa_M\kappa_p}}\right)$$

$$\gamma \le \frac{c_\gamma}{\max\{\kappa_p, \kappa_v\}} \tag{13}$$

where α_2 is given in the appendix and the constants, d_i, are given in the proof below in equation (20). Then, the closed loop system is stochastically stable.

Proof: Consider the following Lyapunov function candidate.

$$V(\tilde{q}, \dot{\tilde{q}}, \tilde{W}) = V_0 + \tfrac{1}{2}\text{Tr}\left[\tilde{W}^T\Gamma^{-1}\tilde{W}\right]$$

where $V_0 = \frac{1}{2}\dot{\tilde{q}}^T K_v M(q)\dot{\tilde{q}} + \frac{1}{2}\tilde{q}^T K_v K_p \tilde{q} + \frac{1}{2}\tilde{q}^T K_p K_v \tilde{q} + \tilde{q}^T K_p M(q)\dot{\tilde{q}}$. It can be shown that there exists positive constants $c_1, c_2, c_3, c_4 > 0$ such that

$$c_1\|\tilde{q}\|^2 + c_2\|\dot{\tilde{q}}\|^2 \leq V_0 \leq c_3\|\tilde{q}\|^2 + c_4\|\dot{\tilde{q}}\|^2$$

We note that the methods employed are similar to that in presented in [15] and [14]. The condition for which $c_1, c_2 > 0$ [13] is given by

$$\kappa_v > \omega_p\omega_M\sqrt{\frac{1}{2\kappa_M\kappa_p}} \quad (14)$$

It can also be shown that

$$\left[\frac{\partial V}{\partial x}\right]^T F(x) \leq -\begin{bmatrix}\|\tilde{q}\|\\ \|\dot{\tilde{q}}\|\end{bmatrix}^T A_0 \begin{bmatrix}\|\tilde{q}\|\\ \|\dot{\tilde{q}}\|\end{bmatrix} + \mathcal{W}(\|\tilde{q}\|, \|\dot{\tilde{q}}\|, K_p K_v)$$

where F and x are defined in (9), and

$$\mathcal{W} = \omega_p\left(\|\dot{\tilde{q}}\|^2\|\tilde{q}\|\xi_2\right) + \omega_v\left(\|\dot{\tilde{q}}\|^3\eta_2\right)$$

Before proceeding, we implement some assumptions that will simplify the forthcoming analysis. First, we set $\beta_p = \beta_v = 1$. These quantities denote the ratio of the maximum and minimum singular values of the feedback gain matrices. In other words, we will take $K_p = \kappa_p I$ and $K_v = \kappa_v I$. Let the learning rate, $\Gamma = \gamma I$. With these assumptions, the matrix A_0 is given as

$$A_0 = \begin{bmatrix} \kappa_p^2 - \kappa_p\alpha_2 & -\frac{\kappa_p(\alpha_3+\xi_1)+\kappa_v\alpha_2}{2} \\ -\frac{\kappa_p(\alpha_3+\xi_1)+\kappa_v\alpha_2}{2} & \kappa_v^2 - \kappa_p\omega_M - \kappa_v(\alpha_3+\eta_1)\end{bmatrix}$$

The constants ξ_i, η_i and α_i are defined in the appendix.

We now compute the Ito anomaly. Using (7), we get

$$\begin{aligned}\frac{1}{2}\text{Tr}\left(V_{xx}GG^T\right) &= \frac{1}{2}\text{Tr}\left(V_{x^{(3)}x^{(3)}}G_3G_3^T\right)\\ &\leq \frac{\gamma}{2}\|\phi\|^2\sigma_B^2\|\tau_{fb}\|^2\end{aligned}$$

where σ_B and ϕ_B were defined previously in (11) and (12), respectively. Let $\gamma \leq \frac{c_\gamma}{\max\{\kappa_p,\kappa_v\}}$.

Then, we have

$$\begin{aligned}&\tfrac{1}{2}\text{Tr}\left(V_{xx}GG^T\right) \leq \tfrac{\phi_B^2\sigma_B^2 c_\gamma}{2\max\{\kappa_p,\kappa_v\}}\|\tau_{fb}\|^2\\ &\leq \tfrac{c_\gamma b_0}{2}\left(\kappa_v\|\dot{\tilde{q}}\|^2 + 2\kappa_p\|\tilde{q}\|\|\dot{\tilde{q}}\| + \kappa_p\|\tilde{q}\|^2\right)\end{aligned}$$

We can now bound LV as follows:

$$LV \leq -\begin{bmatrix}\|\tilde{q}\|\\ \|\dot{\tilde{q}}\|\end{bmatrix}^T A \begin{bmatrix}\|\tilde{q}\|\\ \|\dot{\tilde{q}}\|\end{bmatrix} + \mathcal{W}(\|\tilde{q}\|, \|\dot{\tilde{q}}\|, K_p, K_v) \quad (15)$$

where

$$A = \begin{bmatrix}\kappa_p^2 - \kappa_p(\alpha_2 + \frac{b_0c_\gamma}{2}) & -\frac{\kappa_p f_1+\kappa_v\alpha_2}{2}\\ -\frac{\kappa_p f_1+\kappa_v\alpha_2}{2} & \kappa_v^2 - \kappa_p\omega_M - \kappa_v(\alpha_3+\eta_1+\frac{c_\gamma b0}{2})\end{bmatrix}$$

and $f_1 = \alpha_3 + \xi_1 + c_\gamma b_0$. For the symmetric matrix A to be positive definite, we have the following two conditions: $b^2 < ad$ and $0 < a + d$ where $a = \kappa_p^2 - \kappa_p(\alpha_2 + \frac{b_0c_\gamma}{2})$, $b = -\frac{\kappa_p f_1+\kappa_v\alpha_2}{2}$ and $d = \kappa_v^2 - \kappa_p\omega_M - \kappa_v(\alpha_3 + \eta_1 + \frac{c_\gamma b0}{2})$. The first condition ($b^2 < ad$) yields:

$$\begin{aligned}&\tfrac{1}{4}\left(\kappa_v\alpha_2 + \kappa_p(\alpha_3+\xi_1+c_\gamma b_0)\right)^2 <\\ &(\kappa_p^2 - \kappa_p(\alpha_2 + \tfrac{b_0c_\gamma}{2}))(\kappa_v^2 - \kappa_v(\alpha_3+\eta_1+\tfrac{b_0c_\gamma}{2}) - \kappa_p\omega_M)\end{aligned} \quad (16)$$

The second condition ($a + d > 0$) yields:

$$\begin{aligned}&\kappa_p^2 - \kappa_p(\alpha_2 + \tfrac{b_0c_\gamma}{2})\\ &+\kappa_v^2 - \kappa_v(\alpha_3+\eta_1+\tfrac{b_0c_\gamma}{2}) - \kappa_p\omega_M > 0\end{aligned} \quad (17)$$

Suppose that (16) holds. Then, the left hand side of (19) satisfies

$$\begin{aligned}&\kappa_p^2 - \kappa_p(\alpha_2 + \tfrac{b_0c_\gamma}{2}) + \kappa_v^2 - \kappa_v(\alpha_3+\eta_1+\tfrac{b_0c_\gamma}{2}) - \kappa_p\omega_M\\ &> \kappa_p^2 - \kappa_p(\alpha_2 + \tfrac{b_0c_\gamma}{2}) + \tfrac{1}{4}\tfrac{(\kappa_v\alpha_2+\kappa_p(\alpha_3+\xi_1+c_\gamma b_0))^2}{\kappa_p^2-\kappa_p(\alpha_2+\frac{b_0c_\gamma}{2})}\end{aligned}$$

Setting the right hand side of the above inequality to be positive we get

$$\kappa_p > \alpha_2 + \frac{b_0c_\gamma}{2} \quad (18)$$

Hence, it remains to establish the conditions for which (16) holds. After some manipulation, (16) is equivalent to

$$d_1\kappa_v^2 - d_2\kappa_v - d_3 > 0 \quad (19)$$

where

$$\begin{aligned}d_1 &= 4(\kappa_p^2 - \kappa_p v_3) - \alpha_2^2 & v_1 &= \alpha_3 + \xi_1 + c_\gamma b_0\\ d_2 &= 4v_2(\kappa_p^2 - \kappa_p v_3) + 2\kappa_p\alpha_2 v_1 & v_2 &= \alpha_3 + \eta_1 + \tfrac{b_0c_\gamma}{2}\\ d_3 &= 4(\kappa_p^2 - \kappa_p v_3)\kappa_p\omega_M + v_1^2\kappa_p^2 & v_3 &= \alpha_2 + \tfrac{b_0c_\gamma}{2}\end{aligned} \quad (20)$$

If we require that $d_1, d_2, d_3 > 0$ then we can set κ_v to be greater than right most zero of the left hand side. Setting $d_1 > 0$, we get

$$\kappa_p > \frac{1}{2}\left(v_3 + \sqrt{v_3^2 + \alpha_2^2}\right) \quad (21)$$

Choosing κ_p to satisfy (21) automatically satisfies (18). It also ensures that d_2 and d_3 are positive. Hence, we have

$$\kappa_v > \frac{d_2 + \sqrt{d_2^2 + 4d_1d_3}}{2d_1} \quad (22)$$

The above two inequalities establish the conditions for the feedback gain matrices stated in the theorem. Let $x_1 = \|\tilde{q}\|$, $x_2 = \|\dot{\tilde{q}}\|$ and $x = \left[\begin{smallmatrix}x_1\\ x_2\end{smallmatrix}\right]$. Then,

$$LV \leq A_1\|x\|^3 - A_2\|x\|^2$$

where $A_1 = \frac{\kappa_p\xi_2}{2} + \kappa_v\eta_2$ and $A_2 = \lambda_{min}(A)$. Hence $LV \leq 0$ if $\|x\| < \frac{A_2}{A_1}$. By Theorem 2.2 in [11], it follows that the closed loop system is stable in probability.

■

Remark 1: In the case of SDN on the training signal, the resulting Ito anomaly derived in (15) results in a quadratic term in the performance error. This, in general is not a problem since the derivative of the Lyapunov function is usually bounded by a negative definite quadratic function.

The difficulty arises due to the fact that the coefficients of the quadratic term involve the square of the feedback gain terms. Since this Ito term always increases the energy of the system, we see that increasing the feedback gain will now destabilize the system. This is intuitively clear since increasing the magnitude of the feedback control signal will increase the noise applied to the system. In the deterministic case, there are negative definite terms with coefficients that involve the square of the feedback gain terms. Hence, increasing the feedback gain sufficiently results in a stable system. Thus, in this case, given a fixed bound on the magnitude of the noise intensities, we choose a learning rate that does not amplify the SDN more than the stabilizing feedback control signal. Note also that the Ito term involves the magnitude of the basis function network squared. We note however, that in our case, we use a basis function network that depends only on the desired states. Hence, we can always bound this term for any basis network. On the other hand, if the inputs to the basis network consisted of actual states, then the Ito anomaly would result in a quartic term in the errors complicating the analysis. This illustrates another advantage of using a purely feedforward network.

Remark 2: Note that in the theorem, we simplified the analysis by setting $K_v = \kappa_v I$. This results in $\eta_1 = \eta_2 = 0$. This follows from a skew-symmetric property of the robot arm dynamics and is a common technique employed in the adaptive control of robots [14].

Remark 3: Note the effect of the parameter c_γ on the stability curves. As can be observe from (21), increasing c_γ increases linearly the bound on κ_p. The κ_v bound (22) dependence on c_γ is more complicated but increases as well with increasing c_γ.

V. Simulation Examples

In this section we simulate the 2-link manipulator under the condition of signal dependent noise corrupting the training signal. The 2-link is described by [14].

$$M(q)\ddot{q} + V_m(q,\dot{q})\dot{q} + G(q) = \tau$$

where

$$M(q) = \begin{bmatrix} (m_1+m_2)a_1^2 + m_2 a_2^2 + 2m_2 a_1 a_2 \cos(q_2) & m_2 a_2^2 + m_2 a_1 a_2 \cos q_2 \\ m_2 a_2^2 + m_2 a_1 a_2 \cos q_2 & m_2 a_2^2 \end{bmatrix}$$

$$V_m(q,\dot{q}) = \begin{bmatrix} -\dot{q}_2 m_2 a_1 a_2 \sin q_2 & -(\dot{q}_1+\dot{q}_2) m_2 a_1 a_2 \sin q_2 \\ \dot{q}_1 m_2 a_1 a_2 \sin q_2 & 0 \end{bmatrix}$$

$$G(q) = \begin{bmatrix} (m_1+m_2) g a_1 \cos q_1 + m_2 g a_2 \cos(q_1+q_2) \\ m_2 g a_2 \cos(q_1+q_2) \end{bmatrix}$$

All examples to follow will be with respect to the two-link manipulator dynamics described above with the parameters: $m_1 = 0.8$ kg, $m_2 = 2.3$ kg, $a_1 = 1$ m, $a_2 = 1$ m, and $g = 9.8$ m/s^2.

Example 5.1 (Maximizing the Learning Rate): In Theorem 3., γ is chosen according to (13). To maximize γ, we set

$$\gamma = \frac{c_\gamma}{\max\{\kappa_p, \kappa_v\}}$$

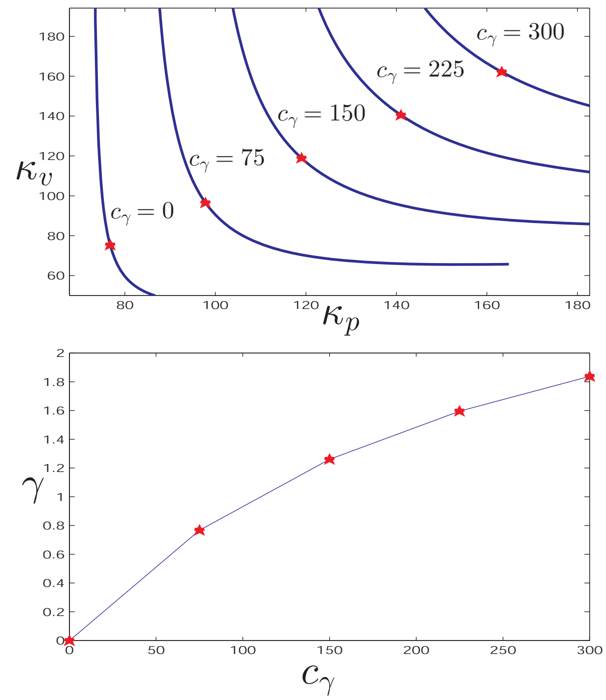

Fig. 2. *Example 5.1: Maximizing the Learning Rate*

and given an arbitrary c_γ, select the feedback gains (which depend on c_γ) such that the denominator is minimized. This occurs when $\kappa_p = \kappa_v$. In Fig. 2, we plot the stability curves for various c_γ. To maximize γ, we choose the points on the stability curves where $\kappa_p = \kappa_v$. This is indicated by a star. In the bottom panel, we plot the corresponding learning rate that maximizes (13). In the simulation, the noise intensity vector is taken to be $\sigma = [\sigma_1\ \sigma_2]^T$ with $\sigma_i = 0.1$, for $i = 1, 2$.

Example 5.2: Instability due to Noisy Teacher: Here we show an example where SDN in the training signal destabilizes the system. We let the noise $\sigma = [\sigma_1\ \sigma_2]^T$ with $\sigma_1 = 0.7$ and $\sigma_2 = 0.5$. We choose the feedback gains according to the nominal stability curve. In this case, we set $\kappa_p = 100$ and $\kappa_v = 120$. The learning rate, γ was set to unity. In Fig. 3., we plot the norm of the combined error, $\|x\| = \sqrt{\|\tilde{q}\|^2 + \|\dot{\tilde{q}}\|^2}$. We observe the resulting destabilizing effect of the noisy teaching signal.

Example 5.3: Stable Learning and Control with SDN: In the previous example, the feedback gains were chosen according to the nominal stability curve which did not take into account the noise on the training signal. In this example, we will show that selecting the feedback gains and learning rate according to Theorem 3 results in stable learning and control of the 2-link manipulator. We take the noise to be $\sigma = [\sigma_1\ \sigma_2]^T$ with $\sigma_i = 0.1$, for $i = 1, 2$. We choose the third point in Fig. 2., which corresponds to $c_\gamma = 150$, and $\kappa_p = \kappa_v = 119.0355$. Using (13), we compute $\gamma = 1.2601$. In Fig. 4., we simulate the teaching signal. In this case stable behavior is observed. In Fig. 4. we also simulate the combined performance error, $\|x\| = \sqrt{\|\tilde{q}\|^2 + \|\dot{\tilde{q}}\|^2}$.

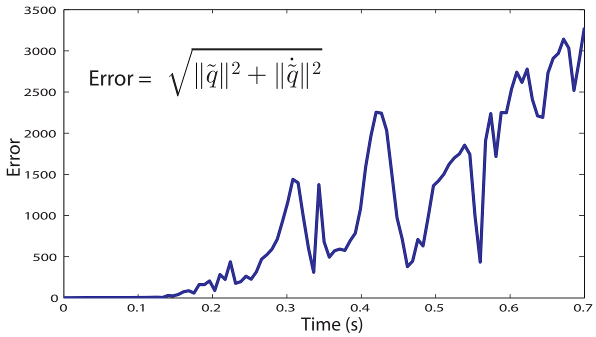

Fig. 3. *Example 5.2: Instability due to Noisy Teacher*

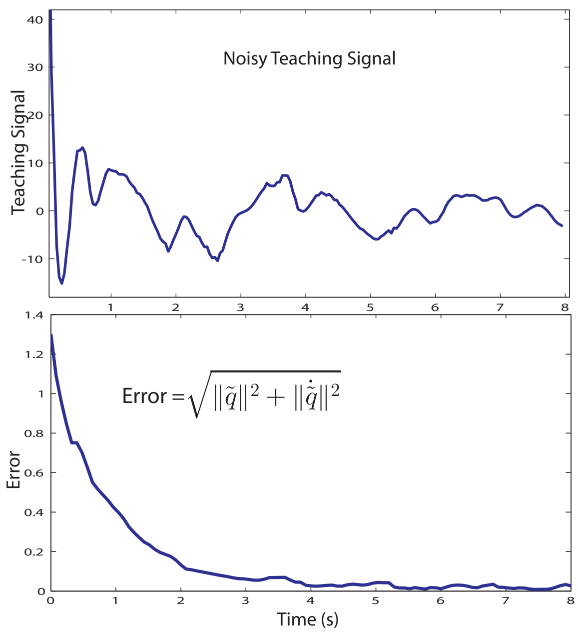

Fig. 4. *Example 5.3: Stable Learning and Control with SDN*

APPENDIX

In this section, we define the bounding constants used in Theorem 1. For a detailed description, see [13]

$$
\begin{aligned}
\xi_1 &= C_V\|\dot{q}_d\| & \alpha_2 &= C_M\|\ddot{q}_d\| + C_{V_q}\|\dot{q}_d\|^2 + C_G \\
\alpha_3 &= C_{V_{\dot{q}}}\|\dot{q}_d\| & C_{V_{\dot{q}}} &= \max_j \sup_q \sum_{i=1}^{n} \|\omega_{ij}(q)\| \\
\eta_1 &= C_V & C_V &= \max_j \sup_q \sum_{i=1}^{n} \|\omega_{ij}(q)\| \\
\omega_{ij}(q) &= [V_i(q)]_{[:,j]} & C_M &= \max_j \sum_{i=1}^{n} \sup_q \left\| \frac{\partial m_{ij}(q)}{\partial q} \right\| \\
& & C_{V_q} &= \max_j \sup_q \sum_{i=1}^{n} \sqrt{\sum_{k=1}^{n} \left\| \frac{\partial \omega_{ij}}{\partial q_k} \right\|^2} \\
& & C_G &= \max_j \sup_q \sum_i \left| \left[\frac{\partial G}{\partial q} \right]_{ij} \right|
\end{aligned}
\tag{23}
$$

REFERENCES

[1] Kawato, M., Furukawa, K., Suzuki, R. (1987). A Hierarchical Neural-Network Model for Control and Learning of Voluntary Movement. *Biological Cybernetics* 57 169-185.

[2] Nakanishi, J., and Schaal, S. (2004). Feedback error learning and nonlinear adaptive control. *Neural Networks*, vol. 16: 1453-1465.

[3] Miyamura, A., Kimura, H. (2002). Stability of feedback error learning scheme. *Systems and Control Letters*, 45(4), 303316.

[4] Garwicz, M. (2002). Spinal reflexes provide motor error signals to cerebellar modules relevance for motor coordination. *Brain Res*. Rev. 40, 152165.

[5] Scott, S.H. (2002). Optimal strategies for movement: success with variability *Nature Neuroscience*. vol.5, 1110-1111.

[6] Sherwood DE, Schmidt RA, and Walter CB. (1988). The force/force-variability relationship under controlled temporal conditions.*J Mot Behav.*. vol.2, 106-16.

[7] Stein RB, Gossen ER, and Jones KE. (2005). Neuronal variability: noise or part of the signal? *Nat Rev Neurosci.*. vol.5, 389-97.

[8] Harris, C. M. and Wolpert, D. M. (1998). Signal-dependent noise determines motor planning *Nature*. vol.394, 780-784.

[9] Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *J. Exp. Psychol.*. vol.47, 381391.

[10] Liao, Y., Fang, S., and Nuttle, H. (2003). Relaxed conditions for radial-basis function networks to be universal approximators, *Neural Networks*, vol. 16, no. 7: 1019-1028.

[11] X. Mao. Stochastic Differential Equations and Applications. Horwood, Chichester, 1997.

[12] M.H. Holmes. Introduction to Perturbation Methods. Springer, 1995.

[13] A.K. Ishihara, J. van Doornik, and T.D. Sanger. Feedback error learning with radial basis function networks. submitted, 2007.

[14] Lewis, F.L., Jagannathan, S. and Yesildirek, A. (1999). *Neural Network Control of Robot Manipulators and Nonlinear Systems*, Taylor and Francis Ltd., London.

[15] Wen, J.T. (1990). A unified perspective on robot control: The energy Lyapunov function approach. *International Journal of Adaptive Control and Signal Processing*, vol. 4. no. 6, pp. 487-500.