

# A NEURAL NETWORK CONTROLLER AUGMENTED TO A HIGH PERFORMANCE LINEAR CONTROLLER AND ITS APPLICATION TO A HDD-TRACK FOLLOWING SERVO SYSTEM

Guido Herrmann <sup>\*,1</sup> Shuzhi Sam Ge <sup>\*\*</sup> Guoxiao Guo <sup>\*\*\*</sup>

<sup>\*</sup> Control and Instrumentation Group, University of Leicester  
University Road, Leicester, LE1 7RH, UK, Tel.: ++44-116-252  
2567, Fax.: ++44-116-252 2619, e-mail: gh17@le.ac.uk

<sup>\*\*</sup> Department of Electrical and Computer Engineering  
National University of Singapore, Singapore 119260

<sup>\*\*\*</sup> A\*-Star Data Storage Institute (DSI)  
5 Engineering Drive 1, Singapore 117608

**Abstract:** The performance of a linear, discrete high performance track following controller in a hard disk drive is improved for its disturbance rejection by augmentation of a discrete non-linear, adaptive neural network (NN) element. The neural network element is deemed to be particularly effective for rejection of bias forces, such as friction. Theoretical and experimental results have been obtained. It is shown theoretically that a NN-element is effective in counteracting a non-linear, system-specific, model-dependent disturbance. The disturbance is assumed to be unknown, with the exception that the disturbance is known to be matched to the plant actuator input range and the disturbance is an (unknown) continuous function of the plant output measurements. In an experiment for a laboratory HDD-servo system, it is shown that the NN-control element improves performance and appears particularly effective for a reasonably small number of NN-nodes.

Copyright ©2005 IFAC

**Keywords:** hard-disk, servo-control, friction, compensation, neural networks

## 1. INTRODUCTION

One significant development has characterized the technological advancement of hard disk drives (HDDs): the increase in track density which has now reached about 100 kTPI (TPI=Track per inch) in commercial products. This development has been driven by consumer demand for higher storage capacity of hard disks and a more varied application area for HDDs which require a smaller form factor (e.g. application in portable equipment such as cameras) or large storage space (e.g. modern 'video' recorders storing digital movies). This increase in data density and in hard disk drive application areas implied the need of significant advancements of the overall HDD-technology: such as highly accurate mechanics for the low-cost Voice Coil Motor (VCM)-actuator in small form-factor drives and a decreased head-disk spacing around 10 nm. Furthermore, the VCM-arm has to be constructed to minimize pivot bearing friction and other bias forces on the arm. Despite the advancement in HDD-technology, these forces are expected to pose a significant problem in future high density HDDs (Gong *et al.*, 2002) for the demanded higher accuracy of the servo-control task, in particular for track following and short span track seeking.

It has been acknowledged that the modelling task for bias forces, in particular for friction, is complex

(Hensen *et al.*, 2003; Oboe *et al.*, 2001; Wang *et al.*, 1994). In actual fact for high precision servo-control systems, it is necessary to investigate friction in smaller dimensions, leading to complex continuous static or dynamic models on the micro-level which depend on both position and velocity of the actuator (Dupont *et al.*, 2002; Wang *et al.*, 2001). One way of dealing with the problems of bias forces is to introduce appropriate servo control techniques which can compensate for these complex but continuous bias forces. Linear techniques such as an increased low frequency controller gain are not sufficient to overcome this non-linear effect, since it has been shown that the non-linear effects, such as friction hysteresis, affect in particular head position changes of about 1  $\mu\text{m}$ . Hence, non-linear methods are necessary for compensation (Gong *et al.*, 2002; Herrmann *et al.*, 2005). Furthermore, the model is not easily obtained for this position-dependent and time-varying bias force. Hence, adaptive neural network compensation appears to be most suitable. The advantages of using neural network techniques for compensation of bias forces have been already tested by Ge *et al.* (1998, pp. 172) in closed loop adaptation for robotic systems and by Huang *et al.* (1998) for HDDs (off-line estimation). Most recently a NN-controller with on-line adaptation in closed loop has been tested by Herrmann *et al.* (2002) and Herrmann *et al.* (2005) for a HDD-servo control system. This controller was particularly developed for manipulators with rigid body dynamics (Ge *et al.*, 1998). The NN-controller structure combining a linear PI(D)-control element with a non-linear adap-

---

<sup>1</sup> This work was completed during a two year Senior Research Fellowship of Guido Herrmann at the A\*-Star, Data Storage Institute, Singapore.

tive NN-control term has shown to be effective for a Voice Coil Motor-actuator in a hard disk drive. Hence, it is here of interest to evaluate the effectiveness of NN-control in combination with a discrete linear, high performance track following controller. Thus, the approach is to design at first a high performance track following controller which achieves a desired bandwidth combined with suitable stability margins. The second step is then to design a NN-controller for bias force compensation. This NN-control element is augmented to the linear high performance controller. This appears to be feasible as non-linear bias force effects are usually observed in low frequency (Abramovitch *et al.*, 1994) which can be compensated by the non-linear NN-element; the linear controller acts in high frequency. This procedure is desirable as a custom-designed industrial linear track following controller can be continued to be used and the non-linear control element acts as an ‘add-on’.

Main principles for the theoretical foundation for this discrete NN-control strategy have been taken from Ge *et al.* (1998) and Wang *et al.* (2001) and in particular from Ge *et al.* (2003). As discussed before, it is assumed that the non-linearity is an (unknown) continuous function of the output measurements: In a HDD, these measurements are the position error signal and an estimate for the velocity. In particular, it is assumed that the bias forces can be modelled as fully matched to the range space of the actuator: Although this assumption is not fully accurate for practical systems, it appears to be suitable for a HDD-servo system, since pivot friction and flex cable forces are bias forces acting likewise the VCM at the pivot of the actuator arm. In particular, Herrmann *et al.* (2005) have shown this approach to be successful. More complex NN-controller structures as discussed by Wang *et al.* (2001) would have to be used in any other case. Hence, this article presents the necessary theoretical background and experimental results showing the effectiveness of the approach for this suggested NN-approach.

## 2. THE NOMINAL CONTROL SYSTEM WITH MATCHED NON-LINEARITY

Consider a linear discrete system with matched unknown non-linearity:

$$\begin{aligned} \mathbf{x}_p(k+1) &= A_p \mathbf{x}_p(k) + \mathbf{b}_p u(k) + \mathbf{b}_p f(\Phi(\mathbf{y}(k))), \\ y(k) &= \mathbf{c}_p \mathbf{x}_p(k) \end{aligned} \quad (1)$$

for  $A_p \in \mathbb{R}^{n \times n}$ ,  $\mathbf{b}_p, \mathbf{c}_p^T \in \mathbb{R}^{n \times 1}$ . The system uncertainty, an unknown non-linearity or system dependent disturbance, is expressed with  $f$  as an unknown function of the vector valued function  $\Phi(\mathbf{y}(k))$ . As  $f$  enters the system via  $\mathbf{b}_p$ , it is *matched* to the range of the actuator. The function  $\Phi(\mathbf{y}(k))$  is known where

$$\Phi = \Phi(\mathbf{y}(k)): R^\psi \rightarrow \mathbb{R}^\phi, \mathbf{y}(k) = [y(k) \ y(k-1) \ \dots \ y(k-\psi+1)]^T$$

and  $\psi, \phi \in \mathbb{N}^+$ . Hence, the function  $\Phi(\mathbf{y}(k))$  may be dependent on the time history of the measurable signal  $y$  but not necessarily on the whole time history.  $\Phi$  is continuous in  $R^\psi$  and remains bounded in a compact subset of  $R^\psi$ , the set for the time history of the output measurement, so that:

$$\|\mathbf{y}(k)\| < K : \|\Phi(\mathbf{y}(k))\| < L, K, L > 0$$

where the Euler norm  $\|\cdot\|$  is used. For reasons of brevity, the abbreviation  $\Phi_k = \Phi(\mathbf{y}(k))$  may be used.

The major problem in this case here is the lack of exact knowledge of  $f$ . However, it is assumed for

the uncertainty  $f$ , that it is a continuous, non-linear function in  $\Phi$ :

$$f: \mathbb{R}^\phi \rightarrow \mathbb{R},$$

and remains bounded for any compact subset in  $\mathbb{R}^\phi$ .

Assume  $u(k) = u_{NL}(k) + u_L(k)$  and there exists a linear controller

$$\begin{aligned} \mathbf{x}_c(k+1) &= A_c \mathbf{x}_c(k) + \mathbf{b}_c y(k), \\ u_L(k) &= \mathbf{c}_c \mathbf{x}_c(k) + d_c y(k), \end{aligned} \quad (2)$$

exponentially stabilizing the plant of (1) for  $u_{NL} = 0$  and  $f = 0$  while achieving ultimate boundedness for  $f \neq 0$ . The control input term  $u_{NL}(k)$  is used later to compensate for the non-linearity  $f_k = f(\Phi(\mathbf{y}(k)))$  more efficiently.

*Remark 1.* For many practical plants, it is always reasonable to assume that the model is strictly proper by introducing a very fast first order pole achieving the strictly proper characteristic. For the high frequency range, it is then assumed that the model is not reliably determined creating some extra, possibly very minor model uncertainty.  $\circ$

The closed loop system is represented by:

$$\begin{bmatrix} \mathbf{x}_p(k+1) \\ \mathbf{x}_c(k+1) \end{bmatrix} = \underbrace{\begin{bmatrix} A_p + \mathbf{b}_p d_c \mathbf{c}_p & \mathbf{b}_p \mathbf{c}_c^T \\ \mathbf{b}_c \mathbf{c}_p & A_c \end{bmatrix}}_{A_g} \underbrace{\begin{bmatrix} \mathbf{x}_p(k) \\ \mathbf{x}_c(k) \end{bmatrix}}_{\mathbf{x}_g(k)} + \underbrace{\begin{bmatrix} \mathbf{b}_p \\ 0 \end{bmatrix}}_{\mathbf{b}_g} (u_{NL}(k) + f_k) \quad (3)$$

where the measurable output of the overall closed loop is:

$$\mathbf{y}_g(k) = \underbrace{\begin{bmatrix} \mathbf{c}_p & 0 \\ 0 & I \end{bmatrix}}_{C_g} \begin{bmatrix} \mathbf{x}_p(k) \\ \mathbf{x}_c(k) \end{bmatrix}.$$

This defines the nominal discrete system  $(A_g, \mathbf{b}_g, C_g)$  for which a non-linear neural network term is designed to compensate for the non-linearity  $f$ .

## 3. DESIGN OF THE NON-LINEAR ADAPTIVE COMPENSATOR

This Section has been inspired by the work by Ge *et al.* (2003) on the discrete NN-control for a special class of non-linear systems. Hence, the approach by Ge *et al.* (2003) has been simplified and suitably adjusted to suit the closed loop system  $(A_g, \mathbf{b}_g, C_g)$  with unknown uncertainty/disturbance  $f$ .

Since  $f$  is continuous in  $\Phi$ , the non-linearity  $f = f(\Phi)$  can be arbitrarily closely modelled (Ge *et al.*, 1998; Wang *et al.*, 2001) via

$$f(\Phi) = w^{*T} s(\Phi) + \epsilon$$

for some neural network basis function vector  $s(\Phi) \in \mathbb{R}^l$  and a respectively large enough number  $l \in \mathbb{N}^+$  of neural network nodes. The vector  $w^*$  is the optimal weight vector while the neural network basis function vector satisfies  $s^T(\Phi) s(\Phi) \leq l$ . The scalar  $\epsilon \in \mathbb{R}$  is the modelling error for which  $|\epsilon|$  usually decreases for an increasing number  $l$  of neural network nodes.

Consider the non-linear controller  $u_{NL}$  as follows:

$$u_{NL} = -\hat{w}^T(k)s(\Phi_k). \quad (4)$$

Hence, the non-linearity is cancelled by an estimate  $\hat{f}(k) = \hat{w}^T(k)s(\Phi_k)$  for which the weight update law is:

$$\hat{w}(k+1) = (1-\sigma)\hat{w}(k) - \Gamma s(\Phi_k)y(k), \quad \sigma > 0 \quad (5)$$

for a positive definite symmetric matrix  $\Gamma \in \mathbb{R}^{l \times l}$ ,  $\Gamma > 0$ , and a forgetting factor  $\sigma$ ,  $0 < \sigma < 1$ . Hence, the estimation error for the weights  $w^*$  is defined by:

$$\tilde{w}(k) = \hat{w}(k) - w^*.$$

Using this approach, stability of the overall control system with augmented NN-controller can be summarized by the following Theorem:

*Theorem 1. There always exist NN-controller parameters  $1 \gg \sigma > 0$  and  $\lambda_{max}(\Gamma)$  small enough and  $\lambda_{min}(\Gamma^{-1})\sigma \gg 0$  large enough so that the linear closed loop control system from (3) augmented with a neural network controller from (4) and a weight update law as from (5) is ultimately bounded stable.  $\diamond$*

Using ideas of Ge *et al.* (2003, Theorem 1 & 2), a sketch of the proof of Theorem 1 is provided in the Appendix discussing also the possibility for optimization of the two adaptation law parameters  $\Gamma$  and  $\sigma$ .

*Remark 2.* As shown in the Appendix (see also Ge *et al.* (2003, Theorem 1 & 2)), persistent excitation of the NN-estimation algorithm, adapting in closed loop, is not necessary. However, it will be discussed for the experiment later that excitation caused indirectly via disturbances can be beneficial.  $\circ$

Note that the parameter choice of a large  $\lambda_{min}(\Gamma^{-1})\sigma$  and a small value for  $\lambda_{max}(\Gamma)$  and  $\sigma$  assures stability but effectively equates to disabling the adaptive NN-controller, achieving nominal linear controller performance only. Hence in this case, the learning algorithm is chosen as slowly acting as possible retaining the NN-weight estimates  $\hat{w}(k)$  close to its initial value  $\hat{w}(0)$ . However, in practice, it is rather of interest to obtain a NN-controller which acts *improving* on the overall control system behaviour. In the next section, the concept of the NN-controller will be used to counteract non-linear bias effects on a linear track following controller of a hard disk drive servo-system.

#### 4. BIAS EFFECTS IN TRACK FOLLOWING OF A HARD-DISK DRIVE SERVO-SYSTEM

In hard disk drives, bias effects are usually caused by bias forces such as forces due to the flex cable or the pivot bearing friction. For application of the NN-approach, it is important that these forces are matched to the range space of the actuator, i.e. the input distribution matrix  $\mathbf{b}_g$ , so that the non-linear control signal  $u_{NL}$  can compensate for them. Although this matching condition is not necessarily satisfied, it may be assumed that a significant part of the bias forces can be regarded as matched. For instance, pivot non-linearities due to friction or bias forces due to the flex cable can be regarded as matched as they (together with the driving force of the VCM) affect the VCM-actuator at the pivot. However, bias forces caused by tribological effects due to head-disk interaction cannot necessarily be regarded as matched. The VCM-actuator is not fully rigid so that the dynamics of the

VCM-arm, the read-write head suspension will always distort the force which is in actual fact then causing the bias force at the pivot due to head-disk interaction. Furthermore, the head position measurement (the position error signal) can be regarded only as an approximated image of the actual pivot position. Thus, it is assumed here that the VCM-actuator is sufficiently stiff to allow the application of the NN-technique.

On a macro-level, friction forces or other bias forces appear to be constant. However, on the micro-level, a bias force requires a much more complex model: Rather complex continuous functions depending on velocity and position of the actuated system are usually employed to model friction characteristics (Wang *et al.*, 2001; Dupont *et al.*, 2002). NN-techniques have shown to be very powerful when modelling these bias forces in dependence on pivot velocity and position measurement (Herrmann *et al.*, 2005): Hence, the model for the bias force  $f = f(\Phi)$  in particular for  $\Phi$  is:

$$\Phi_k = \begin{bmatrix} y(k) \\ y(k) - y(k-1) \end{bmatrix} \approx \begin{bmatrix} y(t) \\ T_s \dot{y}(t) \end{bmatrix}$$

where  $T_s$  is the sampling time of the practical sampled-data control system. As for Wang *et al.* (2001),  $f$  may be modelled using Gaussian radial basis functions for the NN-basis functions of  $s(\Phi_k) = [s_1(\Phi_k) s_2(\Phi_k) \dots s_l(\Phi_k)]^T$ :

$$s(\Phi_k) = e^{-\frac{(y(k)-c_{y/i})^2}{\sigma_{y/i}^2}} \cdot e^{-\frac{(y(k)-y(k-1)-c_{\Delta y/i})^2}{\sigma_{\Delta y/i}^2}}$$

The values of  $\sigma_{y/i}, \sigma_{\Delta y/i} \in \mathbb{R}^+$  and  $c_{y/i}, c_{\Delta y/i} \in \mathbb{R}$ , are the variance and center positions for the position measurement and the velocity estimate of the Gaussian radial basis functions. These parameters need to be chosen by the designer.

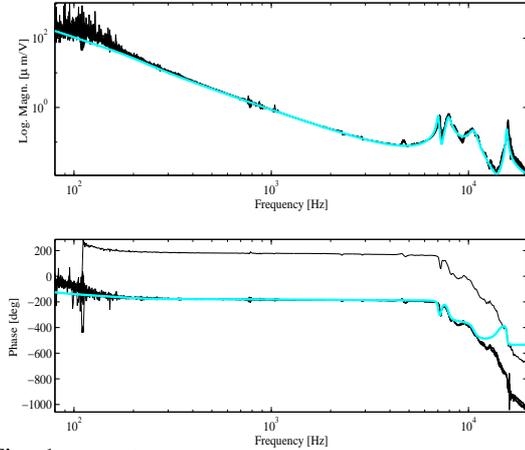


Fig. 1. VCM-frequency responses (including voltage-to-current driver); measured:dark; nominal transfer function: light

#### 5. PRACTICAL IMPLEMENTATION OF THE NN-CONTROL SCHEME

The experiments have been conducted with an up-to-date high performance 3.5 inch VCM-actuator<sup>1</sup> to which a 23 mm suspension<sup>2</sup> with read-write head was attached. The VCM-model measurements in Figure 1 show in particular the effects of friction and other bias

<sup>1</sup> Seagate Technology

<sup>2</sup> Å-Star DSI

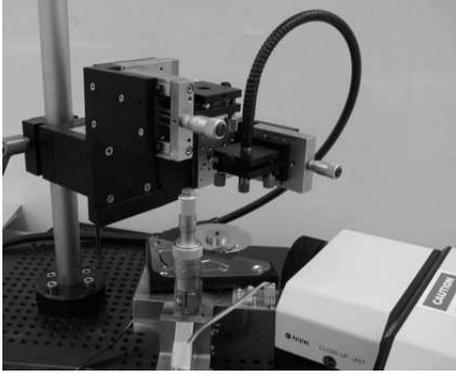


Fig. 2. Experimental Set-up

forces in the frequency range around 100 Hz. Employing a Dynamic Signal Analyser<sup>3</sup> (DSA) and five different excitation amplitudes for ten measurements, the VCM-model measurements show below 100 Hz the typical variation in the frequency response amplitude caused by non-linear effects.

For the practical tests, an A\*-Star DSI designed spin stand was used to mount the read-write head on a commercial 2.5 inch disk rotating at 3300 RPM. The (horizontal) position of the read-write head is measured with a Laser-Doppler-Scanning-Vibrometer<sup>4</sup> (LDV) (Figure 2). From the frequency response measurements for the actuator, we derived from the continuous measurement the linear nominal VCM-actuator models of 12-th order via curve fitting (Figure 1). Using robust linear design techniques, this allowed to design a discrete linear track following controller of 10-th order with an open loop crossover frequency at 1460 Hz and a phase and gain margin of 33° and 4.7 dB. The controller has been implemented with the DSP based system, DS1103<sup>5</sup>, employing a sampling frequency of  $1/T_s = 33$  kHz. Using this nominal linear controller configuration, the NN-controller has been tuned. A description for this follows in the next sections.

### 5.1 Neural network controller design

The characteristics of each NN-node are provided in Table 1. Assuming good knowledge about the practical system, i.e. about the amplitude range of the position and velocity signal during closed loop control, a good preliminary NN-controller is easily obtained. Note that is not necessarily important to fine tune the characteristics of each neural network basis function. It is more important to distribute the NN-nodes over the controller operating region. As the non-linear bias force is unknown, only experimentation can reveal the best choice of NN-nodes. However, it will be seen that the number of required NN-nodes for bias force compensation can be retained small for the investigated case, minimizing the effort for on-line tuning of the NN-controller.

The first step for design is to choose a NN-node with centers  $c_{y/i}$  and  $c_{\Delta y/i}$  directly at the operation point,  $c_{y/i} = 0$ ,  $c_{\Delta y/i} = 0$ , for which the NN-node  $i = 5$  has been selected in Table 1. This node at the operation point is the ‘central’ node which is used for all investigated NN-configurations. As the next step, four nodes were selected which are symmetrically distributed in close proximity of the operation point, i.e. NN-nodes  $i = 2, 3, 7, 8$ . To allow also a good

Table 1. NN-node characteristics in dependence on the node no.  $i$  ( $T_s = 1/33000$ )

$i$	$c_{y/i}$ [ $\mu\text{m}$ ]	$c_{\Delta y/i}$ [ $\mu\text{m}$ ]	$\Gamma_{i,i}$
1	-1	-0.2	$35 \cdot T_s$
2	-0.333	-0.0666	$60 \cdot T_s$
3	0.333	0.0666	$60 \cdot T_s$
4	1	0.2	$35 \cdot T_s$
5	0	0	$200 \cdot T_s$
6	-1	0.2	$35 \cdot T_s$
7	-0.333	0.0666	$60 \cdot T_s$
8	0.333	-0.0666	$60 \cdot T_s$
9	1	-0.2	$35 \cdot T_s$

and accurate model slightly more further off from the operation point, four more NN-nodes  $i = 1, 4, 6, 9$  are to be tested. From these nine nodes, NN-node-configurations are tested which are always symmetric with respect to the central node. The variance values are  $\sigma_{y/i} = \frac{20}{\sqrt{17}}$  and  $\sigma_{\Delta y/i} = \frac{20}{\sqrt{5 \cdot 17}}$  for all nodes to retain a reasonable overlapping of all nine Gaussian radial basis functions. The learning coefficients  $\Gamma = \text{diag}(\Gamma_{1,1}, \Gamma_{2,2}, \Gamma_{3,3}, \dots, \Gamma_{9,9})$  have been chosen as large as possible so that they allow maximal learning speed (Table 1). The forgetting factor  $\sigma = 4 \cdot T_s$  was selected to retain stability of the estimation algorithm but to allow fast adaptation. Using these design characteristics, several combinations of the NN-nodes have been tested for which the results are provided in the next section.

### 5.2 Experimental results

Three different tests have been conducted:

- (1) Error rejection measurement through sine-sweep excitation of the position demand signal using a DSA
- (2) FFT measurement of the position error measurement using a DSA
- (3) Small step responses

All of the three different experiments confirm that the augmented NN-controller can improve track following, in particular, error rejection in the low frequency region is increased: Thus, steady state errors are significantly reduced (see Figure 3(a) versus 3(b)), once the neural network control element is activated. Similar characteristics are observed for the FFT-recordings and the error rejection responses: Consider Figure 4(a) versus Figures 5(a), 6(a), 7(a) and 8(a) in the frequency range up to 100 Hz. The FFT-amplitude is reduced by more than 10 dB when using the NN-element. The appropriate choice of NN-nodes assures an even better result of a 20 dB reduction as seen for Figure 4(a) versus 6(a). Similar results are also observed for the error rejection response in the frequency below 100 Hz as observed for 4(b) versus Figures 5(b), 6(b), 7(b) and 8(b). Improvements of the error rejection of more than 15 dB in the low frequency range are possible.

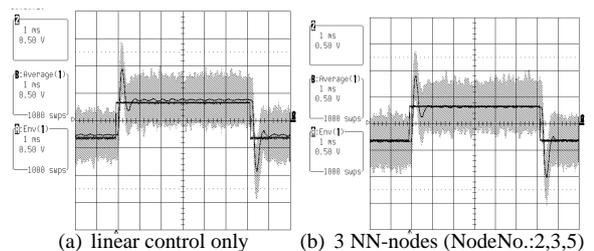


Fig. 3. Step response ( $2\mu\text{m}/\text{V}$ ) (response is averaged; includes envelope of responses (grey shaded)) (physical units: see upper left of each image)

Although for the error rejection responses, the error is always reduced for increasing number  $l$  of NN-nodes, it is observed for the FFT-recordings that the

<sup>3</sup> HP 35670A, Hewlett Packard Company, Washington

<sup>4</sup> Polytec OFV 3001S, Polytec, Waldbronn, Germany

<sup>5</sup> DSpace DS1103 is a product of dSPACE GmbH, Paderborn, Germany

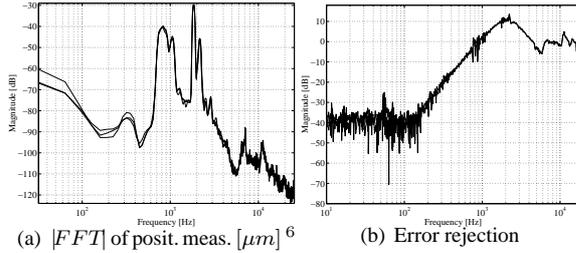


Fig. 4. FFT and error rejection for linear controller only

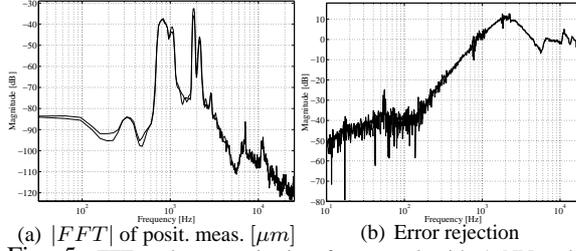


Fig. 5. FFT and error rejection for control with 1 NN-node (NodeNo.:5)

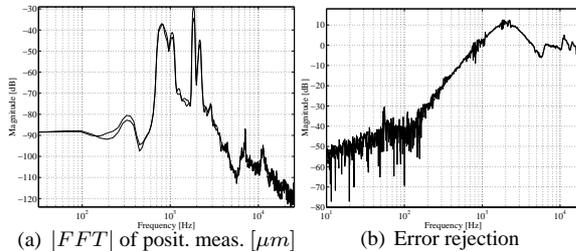


Fig. 6. FFT and error rejection for control with 3 NN-nodes (NodeNo.:2,3,5)

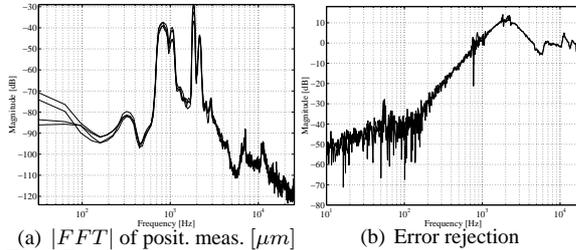


Fig. 7. FFT and error rejection for control with 3 NN-nodes (NodeNo.:1,4,5)

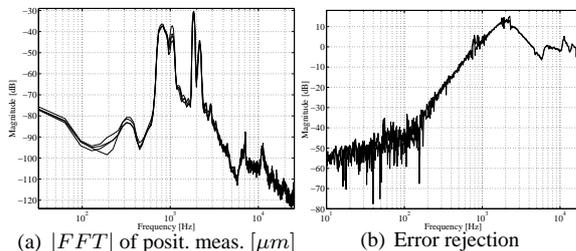


Fig. 8. FFT and error rejection for control with 9 NN-nodes

improvement is dependent on the choice of the NN-node characteristics: Thus, it is important to obtain at first a good model of the non-linear disturbance by selecting the correct NN-node center values close to the operation point (see Figure 5(a) and 6(a)) while center points  $c_{y/i}$  and  $c_{\Delta y/i}$  too far off from the operation point can even act slightly destructive (Figure 6(a) versus 7(a) and 8(a)) (Nevertheless, the NN-controller has been always acting in a positive and improving manner.) Furthermore, it must be also noted that in this case an increase of the NN-node number  $l$  does not necessarily further improve overall performance

as it is obvious from Figure 8 versus Figures 5, 6 and 7. Hence, among those tested configurations, it appears that the neural network control element with three NN-nodes (NodeNo.:2,3,5) is the most suitable.

### 5.3 Discussion

The result appears to indicate that a model of low complexity is sufficient to achieve good disturbance rejection. It is in particular of interest to see, that a NN-control element with the maximum of nine NN-nodes is not able to improve on the result for three NN-nodes (NodeNo.:2,3,5). This seems to point to the problem of *over-learning* or *over-fitting*: For a large number of NN-nodes, the adaptation algorithm attempts to match the given measurements as accurately as possible creating a highly complex model. In case of noise this can lead to the modelling of a wrong representation which does not represent the actual physical disturbance phenomenon. Nevertheless, it has been shown that it is possible to choose the right combination of nodes from a set of nodes well distributed over the operation area. The overall number of available nodes does not need to be large so that it is easily possible to find the ‘optimal’ choice.

As discussed for Theorem 1, persistent excitation of the NN-algorithm is not necessary. However, as for many practically applied adaptive algorithms, it has been noted that external excitation is advantageous. Thus, it has been observed that non-model based disturbances (in particular Non-Repeatable Run-Out disturbances), such as windage, noise, disk flutter, bearing defects etc., appear to improve indirectly the performance of the NN-control element since these disturbances keep the learning process of the NN-controller excited and keep the bias force identification running. Thus, these disturbances can be beneficial to the presented NN-control scheme, continuously learning in closed loop. However, it is certainly not desirable to have these disturbances too large.

## 6. CONCLUSION

A neural network control strategy has been presented which works as augmentation to an existing discrete linear controller. A *theoretical* result has been provided proving stability and performance not worse than the nominal linear controller.

It has been shown *experimentally* for all investigated configurations and any investigated number of NN-nodes, that the *NN-control* technique *improves performance* as it is effective in compensating for low frequency bias effects. These bias forces have been identified as pivot friction and also to a certain extent as head-disk interaction. The neural network controller appears to be especially effective for the low number of three NN-nodes in contrast to the results for nine nodes. A reason for this might be over-fitting and over-learning of the NN-learning algorithm. It has to be investigated if this in general the case for hard disk drives, as this can simplify the tuning of the adaptation algorithm considering a small number of NN-nodes only.

### Appendix A. PROOF OF THEOREM 1

The proof combines linear system theory with ideas from Ge *et al.* (2003, Theorem 1 & 2) on discrete NN-control: Consider a symmetric positive definite Lyapunov matrix:  $P > 0$  for a Lyapunov function

<sup>6</sup> The amplitudes are the result of a root-mean-square averaging process

candidate  $V(k) = \mathbf{x}_g^T(k)P\mathbf{x}_g(k) + \tilde{w}^T(k)\Gamma^{-1}\tilde{w}(k)$  where

$$A_g^T P A_g - P < 0$$

Employing (3), (5) and the relation  $2\tilde{w}(k)^T\Gamma^{-1}\hat{w}(k) = \tilde{w}(k)^T\Gamma^{-1}\tilde{w}(k) + \hat{w}(k)^T\Gamma^{-1}\hat{w}(k) - w^{*T}\Gamma^{-1}w^*$ , the value of  $\Delta V(k) = V(k+1) - V(k)$  can be computed as

$$\Delta V(k) = \begin{bmatrix} \mathbf{x}_g(k) \\ \tilde{w}(k) \\ \hat{w}(k) \end{bmatrix}^T M_1 \begin{bmatrix} \mathbf{x}_g(k) \\ \tilde{w}(k) \\ \hat{w}(k) \end{bmatrix} + \sigma(w^{*T}\Gamma^{-1}w^*) + 2\mathbf{x}_g^T(k)A_g^T P \mathbf{b}_g \epsilon^2 - 2\epsilon \mathbf{b}_g^T P \mathbf{b}_g s(\Phi_k)^T \tilde{w}(k) + \mathbf{b}_g^T P \mathbf{b}_g \epsilon^2 \quad (\text{A.1})$$

where the symmetric matrix  $M_1$  satisfies:

$$M_1 = \begin{bmatrix} M_{11} & -(A_g^T P \mathbf{b}_g + C_g^T) s^T(\Phi_k) & C_g^T \sigma s(\Phi_k)^T \\ * & M_{12} & 0 \\ * & * & (\sigma^2 - \sigma)\Gamma^{-1} \end{bmatrix}, \quad (\text{A.2})$$

$$M_{11} = A_g^T P A_g - P + C_g^T s^T(\Phi_k) \Gamma s(\Phi_k) C_g$$

$$M_{12} = s(\Phi_k) \mathbf{b}_g^T P \mathbf{b}_g s(\Phi_k)^T - \sigma \Gamma^{-1}$$

It is well known that  $\pm 2\mathbf{a}^T \mathbf{b} \leq \mathbf{a}^T \mathbf{a} + \mathbf{b}^T \mathbf{b}$  for two vectors  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^h$ , ( $h \in \mathbb{N}^+$ ). Hence, from this inequality and (A.1), it follows:

$$\Delta V(k) \leq \begin{bmatrix} \mathbf{x}_g(k) \\ \tilde{w}(k) \\ \hat{w}(k) \end{bmatrix}^T M_1 \begin{bmatrix} \mathbf{x}_g(k) \\ \tilde{w}(k) \\ \hat{w}(k) \end{bmatrix} + \sigma w^{*T}\Gamma^{-1}w^* + \delta_1 \mathbf{x}_g^T(k) A_g^T P A_g \mathbf{x}_g + \left( \frac{1}{\delta_1} + \frac{1}{\delta_2} + 1 \right) \mathbf{b}_g^T P \mathbf{b}_g \epsilon^2 + \delta_2 \tilde{w}^T(k) s(\Phi_k) \mathbf{b}_g^T P \mathbf{b}_g s(\Phi_k)^T \tilde{w}(k) = \begin{bmatrix} \mathbf{x}_g(k) \\ \tilde{w}(k) \\ \hat{w}(k) \end{bmatrix}^T M_2 \begin{bmatrix} \mathbf{x}_g(k) \\ \tilde{w}(k) \\ \hat{w}(k) \end{bmatrix} + \sigma w^{*T}\Gamma^{-1}w^* + \frac{\delta_1 \delta_2 + \delta_1 + \delta_2}{\delta_1 \delta_2} \mathbf{b}_g^T P \mathbf{b}_g \epsilon^2 \quad (\text{A.3})$$

where

$$M_2 = \begin{bmatrix} M_{21} & -(A_g^T P \mathbf{b}_g + C_g^T) s^T(\Phi_k) & C_g^T \sigma s(\Phi_k)^T \\ * & M_{22} & 0 \\ * & * & M_{23} \end{bmatrix}$$

$$M_{21} = (1 + \delta_1) A_g^T P A_g - P + C_g^T s^T(\Phi_k) \Gamma s(\Phi_k) C_g$$

$$M_{22} = (1 + \delta_2) s(\Phi_k) \mathbf{b}_g^T P \mathbf{b}_g s(\Phi_k)^T - \sigma \Gamma^{-1}$$

$$M_{23} = (\sigma^2 - \sigma) \Gamma^{-1}. \quad (\text{A.4})$$

As the next step, a zero element  $-\delta_3 V(k) + \delta_3 \mathbf{x}_g^T(k) P \mathbf{x}_g(k) + \delta_3 \tilde{w}(k)^T \Gamma^{-1} \tilde{w}(k)$  is added to the right hand of (A.3) assuming  $\delta_3 < \sigma$ . The terms  $\delta_3 \mathbf{x}_g^T(k) P \mathbf{x}_g(k)$  and  $\delta_3 \tilde{w}(k)^T \Gamma^{-1} \tilde{w}(k)$  are included into  $M_{21}$  and  $M_{22}$  of  $M_2$  so that a matrix  $M_3$  is obtained using  $s^T(\Phi) s(\Phi) \leq l$  and  $\Gamma \leq \lambda_{max}(\Gamma) I$ :

$$\Delta V(k) \leq -\delta_3 V(k) + \begin{bmatrix} \mathbf{x}_g(k) \\ \tilde{w}(k) \\ \hat{w}(k) \end{bmatrix}^T M_3 \begin{bmatrix} \mathbf{x}_g(k) \\ \tilde{w}(k) \\ \hat{w}(k) \end{bmatrix} + \sigma w^{*T}\Gamma^{-1}w^* + \frac{\delta_1 \delta_2 + \delta_1 + \delta_2}{\delta_1 \delta_2} \mathbf{b}_g^T P \mathbf{b}_g \epsilon^2 \quad (\text{A.5})$$

where

$$M_3 = \begin{bmatrix} M_{31} & -(A_g^T P \mathbf{b}_g + C_g^T) s^T(\Phi_k) & C_g^T \sigma s(\Phi_k)^T \\ * & M_{32} & 0 \\ * & * & M_{33} \end{bmatrix}$$

$$M_{31} = (1 + \delta_1) A_g^T P A_g - (1 - \delta_3) P + \lambda_{max}(\Gamma) l C_g^T C_g$$

$$M_{32} = (1 + \delta_2) \mathbf{b}_g^T P \mathbf{b}_g s(\Phi_k)^T + (\delta_3 - \sigma) \Gamma^{-1}$$

$$M_{33} = (\sigma^2 - \sigma) \Gamma^{-1}.$$

Assuming  $M_3 < 0$ , it follows

$$\Delta V(k) \leq -\delta_3 V + \sigma(w^{*T}\Gamma^{-1}w^*) + \frac{\delta_1 \delta_2 + \delta_1 + \delta_2}{\delta_1 \delta_2} \mathbf{b}_g^T P \mathbf{b}_g \epsilon^2$$

which implies ultimate boundedness as the terms  $\sigma(w^{*T}\Gamma^{-1}w^*) + \frac{\delta_1 \delta_2 + \delta_1 + \delta_2}{\delta_1 \delta_2} \mathbf{b}_g^T P \mathbf{b}_g \epsilon^2$  are bounded. Hence, it remains to investigate under what condition  $M_3 < 0$ . From applying Schur's Complement (Boyd *et al.*, 1994), the relationships  $s(\Phi_k)^T s(\Phi_k) < l$ ,  $s(\Phi_k) s(\Phi_k)^T < lI$  and standard matrix upper bounds, it follows that  $M_3 < 0$  if

$$0 > (1 + \delta_1) A_g^T P A_g - (1 - \delta_3) P + \lambda_{max}(\Gamma) l C_g^T C_g + \frac{(A_g^T P \mathbf{b}_g + C_g^T) l (\mathbf{b}_g^T P A_g + C_g)}{(\sigma - \delta_3) \lambda_{min}(\Gamma^{-1}) - (1 + \delta_2) \mathbf{b}_g^T P \mathbf{b}_g l} + \frac{\sigma^2 C_g^T s(\Phi_k)^T s(\Phi_k) C_g}{(\sigma - \sigma^2) \lambda_{min}(\Gamma^{-1})}, \quad 1 > \sigma$$

$$0 > (1 + \delta_2) \mathbf{b}_g^T P \mathbf{b}_g l - (\sigma - \delta_3) \lambda_{min}(\Gamma^{-1}). \quad (\text{A.6})$$

From  $s(\Phi_k)^T s(\Phi_k) \leq l$ , it can be implied that if  $\lambda_{max}(\Gamma)$  and  $\sigma$  are small enough and  $(\lambda_{min}(\Gamma^{-1})\sigma)$  large enough then (A.6) is indeed satisfied and the NN-control system augmented to a linear stabilizing controller is indeed ultimately bounded stable.

*Remark 3.* Although it is usual to tune the adaptation algorithm on-line, the stability proof for the combination of a linear plant with an unknown disturbance and a NN-controller creates the possibility for a non-linear optimization of  $\Gamma$  and  $\sigma$ , improving the overall controller dynamics. In particular, the requirement  $[\mathbf{x}_g^T(k) \tilde{w}^T(k) \hat{w}^T(k)] M_2 [\mathbf{x}_g^T(k) \tilde{w}^T(k) \hat{w}^T(k)]^T + \delta_3 \mathbf{x}_g^T(k) P \mathbf{x}_g(k) + \delta_3 \tilde{w}(k)^T \Gamma^{-1} \tilde{w}(k) < 0$  could be expressed together with  $s(\Phi_k) s(\Phi_k)^T < lI$  in a set of non-convex matrix inequality conditions which could be solved for given  $\delta_3$  to obtain an improved value for  $\Gamma$  and  $\sigma$ . However, this is an area of future research.  $\circ$

## REFERENCES

- Abramovitch, D., F. Wang and G. Franklin (1994). Disk drive pivot nonlinearity modeling part i: Frequency domain. In: *Proc. Amer. Contr. Conf., Baltimore, MD*, pp. 2604–2607.
- Boyd, S., L. El Ghaoui, E. Feron and V. Balakrishnan (1994). *Linear Matrix Inequalities in System and Control Theory*. Soc. Ind. & Appl. Math., Philadelphia.
- Dupont, P., V. Hayward, B. Armstrong and F. Altpeter (2002). Single state elastoplastic friction models. *IEEE Trans. Automat. Contr.* **47**(5), 787–792.
- Ge, S. S., T. H. Lee and C. J. Harris (1998). *Adaptive Neural Network Control of Robotic Manipulators*. World Scientific, Singapore.
- Ge, S. S., T. H. Lee, G. Y. Li and J. Zhang (2003). Adaptive NN-control for a class of discrete-time nonlinear systems. *International Journal of Control* **76**(4), 334–354.
- Gong, J. Q., L. Guo, H. S. Lee and Y. Bin (2002). Modeling and Cancellation of Pivot Nonlinearity in Hard Disk Drives. *IEEE Trans. Magn.*, **38**(5), 3560–3565.
- Hensen, R. H. A., M. J. G. van de Molengraft, and M. Steinbuch (2002). Frequency domain identification of dynamic friction model parameters. *IEEE Trans. Contr. Syst. Technol.*, **10**(2), 191–196.
- Herrmann, G., S. S. Ge and G. Guo (2002). Neural network control of a hard disc drive. In: *Digest Asia-Pacific Magn. Recording Conf.*
- Herrmann, G., S. S. Ge and G. Guo (2005). Practical implementation of a neural network controller in a hard disk drive. *IEEE Trans. Cont. Syst. Techn.* **13**(1), 146–154.
- Huang, T., Y. Ding, S. Weerasooriya and T. S. Low (1998). Disk drive pivot nonlinearity modeling and compensation through fuzzy logic. *IEEE Trans. Magn.* **34**(1), 30–35.
- Oboe, R., A. Beghi, P. Capretta and F. C. Soldavini (2001). Simulator for single stage and dual stage hard disc drives. In: *Proc. IEEE/ASME Intern. Conf. Adv. Intell. Mech.*, pp. 1148–1152.
- Wang, F., T. Hurst, D. Abramovitch and G. Franklin (1994). Disk drive pivot nonlinearity modeling part ii: Time domain. In: *Proc. Amer. Contr. Conf., Baltimore, MD*, pp. 2604–2607.
- Wang, J., S. S. Ge and T. H. Lee (2001). Adaptive friction compensation for servo mechanisms. In: *Adaptive Control of Non-smooth Dynamic Systems* (G. Tao and F. L. Lewis, Eds.), pp. 211–248. Springer, Heidelberg.