Detection of Surface Defects and Servo Signal Restoration for a Compact Disc Player

Peter Fogh Odgaard, *Member, IEEE*, Jakob Stoustrup, *Senior Member, IEEE*, Palle Andersen, and Henrik Fløe Mikkelsen

Abstract—Compact disc (CD) players have been on the market for more than two decades, and the involved technologies, including control are very mature. Some problems, however, still remain with respect to playing CDs having to surface defects like scratches and fingerprints. Two servo control loops are used to keep the optical pick-up unit (OPU) focused and radially locked to the information track of the CD. The problem is to design servo controllers which are well suited for both handling surface defects and disturbances like mechanical shocks. The handling of surface defects requires a low-controller bandwidth which is in conflict with the requirement for the handling of disturbances. This control problem can be solved by the use of a fault tolerant control strategy, where the fault detection is very important. The OPU feeds the controllers with detector signals. Based on these, focus and radial distances can be estimated and a pair of residuals can be computed. This paper introduces a model-based method to estimate focus and radial distances as well as residuals, which have the potential to improve the control of a CD player in case of a surface defect. The methods are applied to experimental data of a scratch. The results of this experimental work show the potential of the method.

Index Terms—Compact disc (CD) players, fault detection, Kalman estimators, servo systems, surface defects.

I. INTRODUCTION

COMPACT DISC (CD) PLAYERS players are widely used today, and have been on the market for more than two decades. However, even though CD players have been on the market for such a long period, performance issues are still to be improved on. It is common to have problems with CDs with surface defects like scratches, fingerprints, etc. These defects can cause the CD player to jump randomly to another track, or in some cases to stop playing the disc. An optical pick-up unit (OPU) is used to retrieve the data/music saved on the disc. The OPU has no physical contact with the disc. Instead two servo loops are formed to keep the OPU focused at the disc's reflection layer and radially tracked over the spiral shaped data track. The OPU includes optics which facilitate indirect measures of the focus and radial distances, each calculated as the difference of signals from two photodiodes. The distances are the distance

P. F. Odgaard, J. Stoustrup, and P. Andersen are with the Department of Control Engineering, Aalborg University, Fredrik Bajers Vej 7C, DK-9220 Aalborg East, Denmark.

H. F. Mikkelsen is with the Bang & Olufsen a/s, Peter Bangs Vej 15, DK-7600 Struer, Denmark.

from the actual position to the position where the OPU is focused and radially tracked. These focus and radial distances are also important in terms of detecting surface defects. The real problem with surface defects is that they degenerate the indirect measures of focus and radial distances, whereas missing information can be reconstructed due to redundancy in data. A possible method used to handle these defects is to use a fault tolerant control strategy, where the surface defects are considered as faults which have to be handled in a special way. A fault is detected as fast as possible and when a fault is detected, the control strategy is changed in a way that accommodates the detected fault. A frequently used method for detection of faults is to observe changes in either the sum of focus signals or the sum of the radial signals, since a fault would reduce these sums, see [1], [2], and [3]. One method used to handle such faults is not to rely on sensor information during the fault. The sensor signals are simply fixed to zero during the fault. The use of the sum signals for detection can be a nonoptimal-detection strategy due to the cross-couplings between distances and detector signals, for example a change in the radial distance also changes the focus sum.

The research in control of CD players has been intense in other directions than fault tolerant control, especially in adaptive and robust controllers applied to the CD player. The first application of a μ-controller used in a CD player was reported in [4], which was based on DK-iterations. An example of an adaptive control design was [5] where a self-tuning controller was suggested. Automatic adjustment of gain in dependence of the reflective characteristics is standard in commercial CD players. A large number of different control strategies were applied to the CD player. An adaptive repetitive method was suggested in [6], quantitative feedback theory was used in [7]. In [8] rejection of nonrepeatable disturbances was used, fuzzy control was used by [9]. In [10], a hybrid fuzzy control was designed, a linear quadratic Gaussian control was used in [11], and disturbance observer was designed in [12]. A vibration absorber to damp the mechanical disturbances was used in [13].

The method suggested in this paper is based on a model of the optical detector system. This optical detector system is modeled as a static mapping from focus and radial distance to the detector signals, see [14]. The use of this mapping has the potential to improve the estimates of focus and radial distances and to calculate residuals useful for detections of faults. A large amount of work has been performed in modeling and identification of the mechanical and electromagnetic parts of CD players. The modeling of these parts of the system is described in [15] and [16]. Vidal (see [16]) describes a simple method to perform
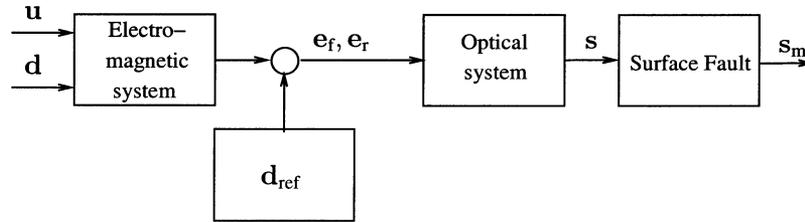
Fig. 1. The principles of the model of the CD player. The CD player consists of four parts. The electromagnetic system and the optical system, the unknown reference $\mathbf{d}_{\mathrm{ref}}[n]$ and the surface fault. $\mathbf{u}[n]$ is the control signal to the electromagnetic system, $\mathbf{d}[n]$ is the unknown disturbances to the electrical-magnetic system, $\mathbf{d}_{\mathrm{ref}}[n]$ is the unknown reference to the system, $e_{\mathrm{f}}[n]$ is the focus distance, $e_{\mathrm{r}}[n]$ is the radial distance, $\mathbf{s}_{\mathrm{m}}[n]$ is the measured detector signals, and $\mathbf{s}[n]$ is detector signals without surface faults.

open loop system identification. Both Yeh (see [17]) and Dettori (see [18]) performed some work on the cross-couplings in the mechanical and electromagnetic parts between focus and radial loop. Regarding the optical part of the system the present control strategies are based on simple linear models not considering the optical cross-coupling [15] and [19], although some work has been done regarding the optical model. Some considerations about the optical cross-couplings are done in [17]. Vidal (see [3] and [20]) deals with models of the optical signal of the focus error without consideration about the cross-coupling with the radial loop. However, a model of the optical system is developed in [14], for the three beam single Foucault detector system.

The main contributions of this paper are: estimation of pick-up positions during surface faults. This estimation also indicate gives a set of residuals indicating surface faults. These residuals are also derived in this paper, and they are decoupled from the focus and radial distances, and an estimate of the focus and radial distances which are valid during surface faults can be found. In order to achieve these estimates, a model of the optical detector is presented, as well as a model of the surface faults. A method to find the inverse map of the combined model of the optical detectors and surface faults is given, as well as a Kalman estimator designed to estimate the focus and radial distances.

The idea of this paper is first to use the optical model to find a local inverse map of this model, and thereby a static estimate of the distances based on the detector signals. This is accomplished by using a method based on a model of the surface faults and Newton-Raphson's method [21]. Based on the model of the faults it is possible to identify two fault parameters which are decoupled from the distances, and which can be used for fault detection. The idea behind these fault parameters can be found in [22]. A Kalman estimator with an internal reference model describing variations in the track is designed to dynamically estimate focus and radial distance for control purposes, see [22]. Finally, these methods are used on experimental data, and the fault detection properties of the fault parameters are tested with a simple threshold detection.

II. SYSTEM DESCRIPTION

This paper is based on a given model structure of the CD player which is shown in Fig. 1. This model structure is based on model structures in [23], [3], and [19]. The first part of the model is the electromagnetic system, which moves the OPU in focus and radial directions. $\mathbf{u}(t)$ is the control signals to the system. $\mathbf{d}(t)$ is disturbances such as harmonic shocks. The outputs of this system are absolute focus and radial distances. Adding to these $\mathbf{d}_{\mathrm{ref}}(t)$, the unknown reference to the system, gives relative focus and radial distances $e_{\mathrm{f}}(t)$ and $e_{\mathrm{r}}(t)$. These distances are sensed by the optical system, which maps the distances to four detector signals, where, traditionally, two of them are used for focus control and the last two are used for radial control. The last part of the model is a model of possible faults. In case of a fault, the detector signals will be degenerated.

In the normal servo system indirect measures of focus and radial distances are used. These indirect measures of the distances are found by taking differences between the two focus and between the two radial signals. It is important to distinguish between disturbances, which the controller shall reject and faults which the controller shall not react to, but using the indirect measures can make the separation of disturbances from faults more difficult. There are two problems with the indirect measures. The first problem is that there are cross-couplings between the focus distance and the radial detector signals and between the radial distance and focus detector signals. The second problem is that the mappings from the distances to the indirect measures are only linear for small distances. The third problem is that the surface faults will also influence on these optical measurements. All these problems lead to the fact that it would be better to establish estimates of focus and radial distances for the controlling of the focus and radial distances, by computing the inverse map of the combined optical and surface fault models. The optical model is a mapping from focus and radial distance to the four detector signals. This mapping is well defined, but is not globally invertible, see [14]. The inverse map is found based on an iterative method using Newton-Raphson's method. This method uses that the variations in focus and radial distances from sample to sample are limited due to the low-pass effect of the suspension of the OPU. This makes it possible to use a locally defined inverse map, and give estimates of focus and radial distances, $\tilde{e}_{\mathrm{f}}[n]$ and $\tilde{e}_{\mathrm{r}}[n]$ at each sample n.

However, there is a stochastic part in each fault. These signal parts together with measurement noise add some noise to $\tilde{e}_{\mathrm{f}}[n]$ and $\tilde{e}_{\mathrm{r}}[n]$, but this noise can be reduced by the use of a Kalman estimator with an internal reference model, see the system structure in Fig. 2. The internal reference model is a model of the eccentricity and skewness of the disc. The Kalman estimates $\hat{e}_{\mathrm{f}}[n]$ and $\hat{e}_{\mathrm{r}}[n]$ can, for example, be used as partial state estimates to a controller, or be fed into the optical model, with the purpose of giving an estimate $\hat{\mathbf{s}}[n]$ of the detected signals $\mathbf{s}[n]$ without
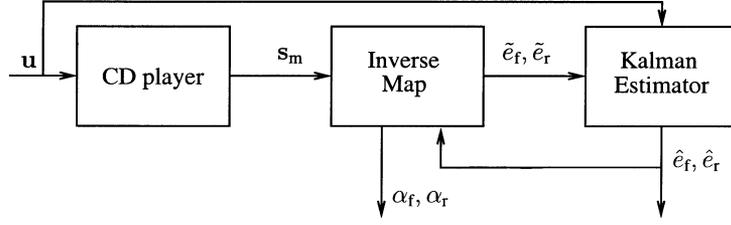
Fig. 2. The structure of the method described in this paper. The method consists of two parts, which estimates fault parameters, (α_f, α_r), focus distance and radial distance from the detector signals. The parts of the method are: inverse map and Kalman estimator. $\mathbf{u}$ is the control signals to the CD player, see Fig. 1, $\mathbf{s}_m$ is the measured detector signals, $\bar{e}_f, \bar{e}_r$ is the static estimated focus and radial distances, and $\hat{\mathbf{e}}_f, \hat{\mathbf{e}}_r$ is the dynamical estimated focus and radial distances.

faults. This dynamical estimate of the detector signals can together with $\tilde{e}_f[n]$ and $\tilde{e}_r[n]$ be used for extraction of the fault parameters.

In this paper, a method is suggested for extraction and estimation of focus and radial distances and fault parameters/residuals. These signals can be used to improve the performance of the nominal and the fault tolerant controllers for a CD player. The system is illustrated in Fig. 2. The symbols in the figure and the rest of the paper are defined as follows: $\mathbf{s}_m[n]$ is a sample of detector signals, $\tilde{e}_f[n], \tilde{e}_r[n]$ are static estimates of focus and radial distances, $\hat{e}_f[n], \hat{e}_r[n]$ are dynamical estimates of focus and radial distances. This means that only two variables are estimated based on the four detector signals. The two remaining dimensions are parameterized by the two fault parameters $\alpha_f[n], \alpha_r[n]$, which parameterize the fault model. These fault parameters are close to zero in case of no faults and large in case of a fault. A more detailed description of the fault model is given in Section III-C. The inverse map solver finds static estimates of optical distances from sampled detector signals.

The four detector signals $D_1[n], D_2[n], S_1[n], S_2[n]$ can be described as follows:

$$\begin{bmatrix} D_1[n] \\ D_2[n] \\ S_1[n] \\ S_2[n] \end{bmatrix} = \mathbf{f}_e(e_f[n], e_r[n], \alpha_f[n], \alpha_r[n]) \quad (1)$$

and

$$\mathbf{f}_e(e_f[n], e_r[n], 0, 0) = \mathbf{f}(e_f[n], e_r[n]). \quad (2)$$

$\mathbf{f}_e(\cdot)$ is an extended model combining the optical and a fault model. $\mathbf{f}(\cdot)$ is the optical model, which maps from the two-dimensional (2-D) error signals into a 4-D space.

The definition of $\mathbf{f}_e(\cdot)$ and $\mathbf{f}(\cdot)$ makes a definition of $\mathbf{s}[n]$ and $\mathbf{s}_m[n]$ possible

$$\mathbf{s}_m[n] = \mathbf{f}_e(e_f[n], e_r[n], \alpha_f[n], \alpha_r[n]) \quad (3)$$

and

$$\mathbf{s}[n] = \mathbf{f}(e_f[n], e_r[n]). \quad (4)$$

Combining the model and the four detector signals gives four equations with 2 variables concerning the computation of the two distances. In order to describe the 4-D space two extra variables are used to describe the deviation from the two dimensional manifold described by (4). These variables are the implicit fault model, parameterized by $\alpha_f[n]$ and $\alpha_r[n]$. This
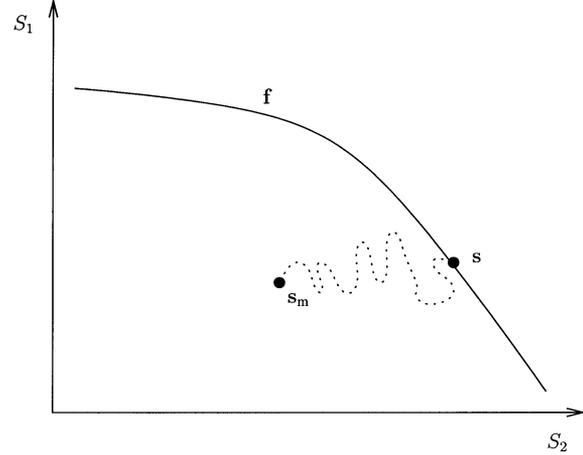

Fig. 3. Due to a fault the measured detector signal, $\mathbf{s}_m[n]$, lies outside the output set of the optical model, $\mathbf{f}(\cdot)$. This is illustrated with the detector signals S_1 and S_2, where $e_f = 0$ implying that $\mathbf{f}(\cdot)$ is a curve in the S_1-S_2 plane and constant in D_1 and D_2.

means that the solution to this problem is to find the inverse map of $\mathbf{f}_e(\cdot)$. The inverse is not globally defined, meaning that the solver only finds a local solution.

In [21], the following definition is stated.

1 Definition (The Disturbance Set): The disturbance set $\mathcal{D} \in \mathcal{R}^4$ is defined as the set in which any sample $\mathbf{s}_m \in \mathcal{R}^4$ of the detector signals, will be located if only disturbances occur, (and no faults). This means that

$$\mathcal{D} = \{\mathbf{s}_m \in \mathcal{R}^4 : \exists e_f, e_r \text{ such that } \{\mathbf{s}_m \leftarrow (e_f, e_r, 0, 0)\} = \mathbf{Im}(\mathbf{f})$$

This definition can be explained in the way, that if no fault occurs the sampled detector signals will be in the disturbance set $\mathcal{D}$, which is a 2-D manifold in $\mathcal{R}^4$. It also means that if sampled detector signals are outside $\mathcal{D}$, this variation is due to a surface fault.

The extended model output for two detector signals is illustrated in Fig. 3. This figure illustrates a sampling of S_1 and S_2, outside $\mathcal{D}$ due to a fault in a case where D_1 and D_2 are supposed to have the nominal value.

The exposition above focuses on discriminating disturbances and faults and, thus, disregards measurement noise. In practice, however, even for a high quality disc, small deviations from this idealized situation will appear, due to small irregularities and electrical noise.

III. MODEL OF THE CD PLAYER

In Fig. 1 the model structure of the CD player is illustrated. It is seen that the model consists of four submodels which all are described in this section. The submodels are described in the following order: a model of the electromagnetic system, a model of the reference signal, a model of the optical system, a fault model, and a reference model.

A. Model of the Electromagnetic System

The OPU includes a 2-axis device, enabling a movement of the OPU vertically for focus correction and horizontally for radial correction of focus and radial errors. Linear electromagnetic actuators are used for both the focus corrections and the radial corrections. The electromagnetic force in the actuators is controlled by focus and radial control voltages $u_f(t)$ and $u_r(t)$. The OPU itself can be modeled as a mass-spring-damper system, with one or two masses dependent on the required details. This results in a second or fourth order model for both focus and radial actuator dynamics, see [19], [16], and [15]. In [16] a system identification on the same CD player setup, as used for experimental work in this paper, is performed. The second order model found in [16] and [3] will be used in this paper. Focus and radial models are of the following structure

$$\dot{\boldsymbol{\eta}}(t) = \begin{bmatrix} -a_0 & -a_1 \\ 1 & 0 \end{bmatrix} \cdot \boldsymbol{\eta}(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \cdot u(t) \qquad (5)$$

$$x(t) = \begin{bmatrix} 0 & b \end{bmatrix} \cdot \boldsymbol{\eta}(t) \qquad (6)$$

where $\eta(t)$ is the state vector, $x(t)$ is the output vector, which is absolute focus or radial position, $u(t)$ is the input vector, which is the control voltage. a_0, a_1, b are model parameters. The values of a_0, a_1, b are found in [16] and [3].

1) Internal Reference Model: The reference signals to focus and radial loops are unknown. The nature of these references is eccentricity and skewness of disc etc. However, it is known that the first and most dominant harmonic of the unknown reference is the angular velocity of the disc. The interval of the angular velocity can be retrieved from [19] to be 4–9 Hz. One way to model the reference is by a bandpass filter with the bandpass region between 4 and 9 Hz. The model will be of the structure

$$\dot{\boldsymbol{\eta}}_{\text{ref}}(t) = \mathbf{A}_{\text{ref}} \cdot \boldsymbol{\eta}_{\text{ref}}(t) + \mathbf{E}_{\text{ref}} \cdot q_{\text{ref}}(t) \qquad (7)$$

$$d_{\text{ref}}(t) = \mathbf{C}_{\text{ref}} \cdot \boldsymbol{\eta}_{\text{ref}}(t) \qquad (8)$$

where: $\eta_{\text{ref}}(t)$ is the state vector of the internal reference model, $d_{\text{ref}}(t)$ is the reference, $q_{\text{ref}}(t)$ is the input to internal reference model which is a white noise signal.

Now it is possible to merge the internal reference model together with focus and radial models. Starting with defining the state vector of the merged CD player model $\boldsymbol{\eta}_{\text{CD}}(t)$.

$$\boldsymbol{\eta}_{\text{CD}}(t) = \begin{bmatrix} \boldsymbol{\eta}_f(t) \\ \boldsymbol{\eta}_{r,f}(t) \\ \boldsymbol{\eta}_r(t) \\ \boldsymbol{\eta}_{r,r}(t) \end{bmatrix} \qquad (9)$$

where: $\boldsymbol{\eta}_f(t)$ and $\boldsymbol{\eta}_r(t)$ are the respective focus and radial state vectors. $\boldsymbol{\eta}_{r,f}(t)$ and $\boldsymbol{\eta}_{r,r}(t)$ are the respective focus and radial internal reference state vectors.

The control signal vector is

$$\mathbf{u}_{\text{CD}}(t) = \begin{bmatrix} \mathbf{u}_f(t) \\ \mathbf{u}_r(t) \end{bmatrix} \qquad (10)$$

where $\mathbf{u}_f(t)$ and $\mathbf{u}_r(t)$ are, respectively, the focus and radial control signals.

The input to both the internal reference models are two noise signals $q_{\text{ref},f}(t)$ and $q_{\text{ref},r}(t)$, and are merged as column elements in the vector $\mathbf{q}_{\text{ref}}(t)$. The focus and radial models are given in (5) and (6), and the focus and radial internal reference models are given in (7) and (8).

The focus and radial actuators are decoupled, and the focus model and the focus internal reference model are connected by the outputs: $e_f(t) = x_f(t) + d_{\text{ref},f}(t)$ and the radial distance is $e_r(t) = x_r(t) + d_{\text{ref},r}(t)$. $x_f(t)$ and $x_r(t)$ denote, respectively, focus and radial positions, and $d_f(t)$ and $d_r(t)$ denote, respectively, focus and radial references. Based on these definitions it is possible to state the extended CD player model of the focus and radial actuators and the focus and radial internal reference models.

$$\dot{\boldsymbol{\eta}}_{\text{CD}}(t) = \mathbf{A}_{\text{CD}} \cdot \boldsymbol{\eta}_{\text{CD}}(t) + \mathbf{B}_{\text{CD}} \cdot \mathbf{u}_{\text{CD}}(t)$$
$$+ \mathbf{E} \cdot \mathbf{q}_{\text{ref}}(t) \qquad (11)$$

$$\begin{bmatrix} e_f(t) \\ e_r(t) \end{bmatrix} = \mathbf{C}_{\text{CD}} \cdot \boldsymbol{\eta}_{\text{CD}}(t) \qquad (12)$$

where

$$\mathbf{A}_{\text{CD}} = \begin{bmatrix} \mathbf{A}_f & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{r,f} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_{r,r} \end{bmatrix} \in \mathcal{R}^{8 \times 8} \qquad (13)$$

$$\mathbf{B}_{\text{CD}} = \begin{bmatrix} \mathbf{B}_f & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_r, \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \in \mathcal{R}^{8 \times 2} \qquad (14)$$

$$\mathbf{C}_{\text{CD}} = \begin{bmatrix} \mathbf{C}_f & \mathbf{C}_{r,f} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{C}_r & \mathbf{C}_{r,r} \end{bmatrix} \in \mathcal{R}^{2 \times 8} \qquad (15)$$

$$\mathbf{E} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{E}_{r,f} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \\ \mathbf{0} & \mathbf{E}_{r,r} \end{bmatrix} \in \mathcal{R}^{8 \times 2} \qquad (16)$$

where $\mathbf{A}_f, \mathbf{B}_f, \mathbf{C}_f$ are the model matrices in the focus model, and $\mathbf{A}_r, \mathbf{B}_r, \mathbf{C}_r$ are the model matrices in the radial model. $\mathbf{A}_{r,f}, \mathbf{C}_{r,f}, \mathbf{E}_{r,f}$ are the matrices of the focus reference model, and $\mathbf{A}_{r,r}, \mathbf{C}_{r,r}, \mathbf{E}_{r,r}$ are the matrices of the radial reference model.

B. Optical Model

One of the most important things, in making a CD player work, is to focus the laser beam at the information track. The CD player used in this work has its detector system implemented as a hologram, but works as a three beam single Foucault detector system. The main beam is used both to retrieve the information saved on the disc, and to focus the beam at the disc reflection layer. Two additional beams are used to keep the main beam radially tracked. The information in the track is stored by using two different levels, called pit and land, (the land has the same
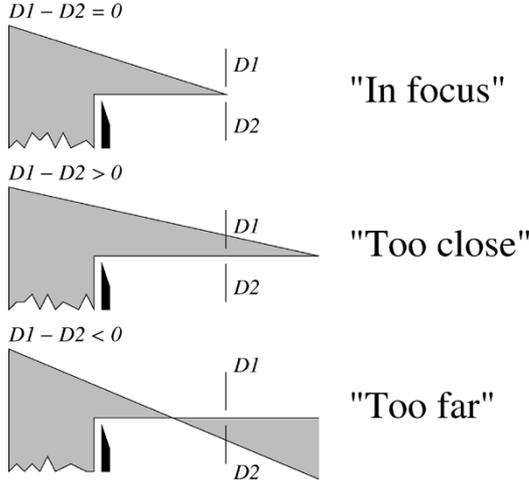
Fig. 4. The three possible situation of the single Foucault focus detector principle. In the first situation the spot is focused at the detectors. The second situation shows a spot focused behind the detector. The third situation shows a spot focused in front of the detectors.

level as the area outside the track). The level difference is a quarter of the wave-length of the laser beam in the material. That is when the light is reflected, the light reflected from a pit interferes destructively with the light reflected from a land, see [15] and [24]. These pits and lands are also important in the task of modeling the optics of the system. In the following, focus and radial detector principles are shortly described.

1) The Focus Detector: The focus detector consists of two detectors, $D_1[n]$ and $D_2[n]$. The idea is to introduce some asymmetry in the light path from the disc surface to these detector in such a way, that

> if the light beam is focused : $D_1[n] - D_2[n] = 0$,
> if the OPU is too far away : $D_1[n] - D_2[n] > 0$,
> if the OPU is too close : $D_1[n] - D_2[n] < 0$.

This asymmetry can be generated in a number of ways, for example by the single Foucault principle which is illustrated in Fig. 4. The idea behind this focusing principle is to place a knife edge into the light path, such that only half of the light beam passes the knife edge, and the rest is absorbed by the knife edge. The optical system is designed such that if the light beam is focused on the disc it would also be focused on the detectors. In the cases where the OPU is either too close to or too far away from the disc, the beam focus point would be either behind or in front of the detectors. The principle illustrated in Fig. 4 is based on the assumption that the light emits from a point source. However, the light spot on the disc is disc shaped with distribution of energy shaped as a sinc function. In the following this light energy distribution is simplified. A more detailed description is given in [14]. In the simplified model the energy distribution of the light source is assumed to be uniform [14]. This is handled by convolving the model output with a sinc shaped signal.

2) The Radial Detector: Fig. 5 illustrates how the three beams are placed relative to each other on the disc surface. The main spot in the middle and the two others are placed one to each side of the track, with a distance from their center to the center of the main spot a_k. If the pick-up is located symmetrically over the track, the two satellite spots will cover
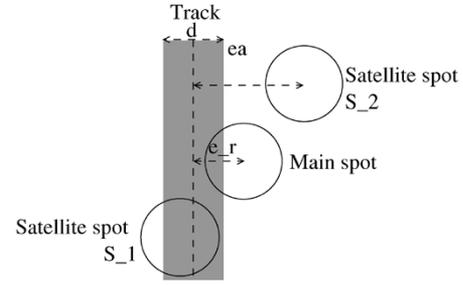

Fig. 5. Illustration on how the three beams are placed relative to each other on the disc surface.

equal areas of the track. In the detector signals only the temporal mean of the pit and land is seen, simply due to the fact that the disc spins around and that the detector signals are low-pass filtered. Fig. 5 illustrates a situation where the pick-up is too much to the right. In this case, $S_2[n]$ receives more light energy than $S_1[n]$, due to the destructive interferences, that is $S_1[n] - S_2[n] < 0$. On the other hand, if the pick-up is too much to the left $S_1[n] - S_2[n] > 0$.

3) The Focus and Radial Optical Models: This optical model is expressed by the mappings, described in (17)–(20).

$$f_1 : (e_f[n], e_r[n]) \rightarrow D_1[n] \tag{17}$$

$$f_2 : (e_f[n], e_r[n]) \rightarrow D_2[n] \tag{18}$$

$$f_3 : (e_f[n], e_r[n]) \rightarrow S_1[n] \tag{19}$$

$$f_4 : (e_f[n], e_r[n]) \rightarrow S_2[n]. \tag{20}$$

These mappings can be simplified using the following simplification

$$f_i(e_f[n], e_r[n]) \approx h_i(e_f[n]) \cdot g_i(e_r[n]) \tag{21}$$

where

$$i \in \{1, 2, 3, 4\} \tag{22}$$

Moreover

$$g_1(e_r[n]) = g_2(e_r[n]). \tag{23}$$

This simplification only has a small influence, the maximum relative variations found in simulations are 6% for e_f in the interval $[-0.3 \, \mu m \cdots 0.3 \, \mu m]$.

In the present experimental setup it is only possible to measure either a varying $e_f[n]$ with a constant $e_r[n]$ or a varying $e_r[n]$ with a constant $e_f[n]$. This means that it is not possible to experimental validate (21) completely in experiments.

Equation (21) can be interpreted as follows. The $g_i(e_r[n])$ function computes the maximal detected energy for a given value of $e_r[n]$. $h_i(e_f[n])$ computes how large a ratio of the reflected energy in the spot which is detected.

In [14], detailed optical models are described. These models are unfortunately not differentiable at every point. However, an even more realistic model would be differentiable at every point, meaning that it is the model simplifications which introduce these non differentiable points. In [14] and [21] it is suggested to approximate $h_i(e_f[n])$ and $g_i(e_r[n])$ with cubic splines. Due to this not-differentiable-at-all-points property of the developed
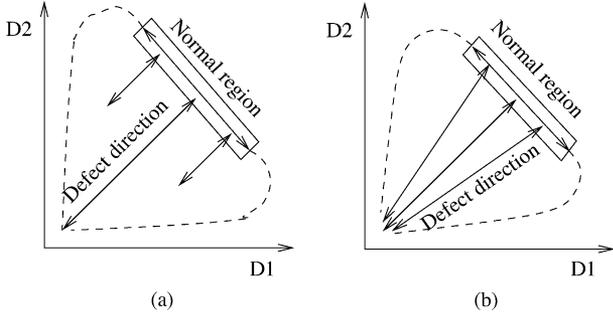
Fig. 6. A 2-D illustration of the two fault models. Model A is the orthogonal fault model, and Model B is the scaling fault model. The fault moves the detector signals outside the normal operation region. The dashed line illustrates detector signal path outside the normal region.
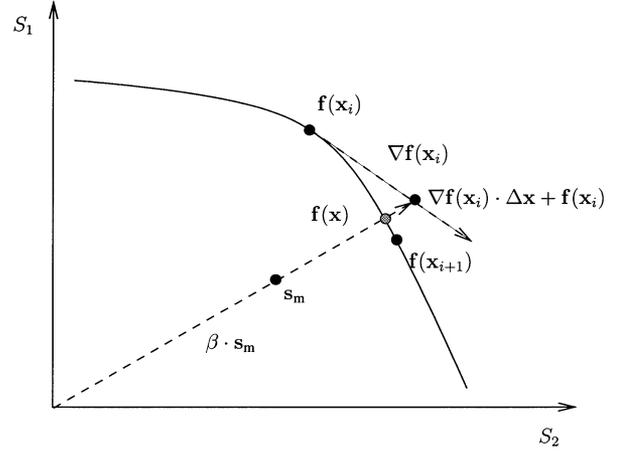

Fig. 7. Illustration of the principles of the scaling projection method. The illustration shows an example in $\mathcal{R}^2$. $\mathbf{f}(\mathbf{x}_i)$ is the starting point for the iteration, $\nabla \mathbf{f}(\mathbf{x}_i)$ is the gradient at the starting point, $\mathbf{s}_m$, $\beta \cdot \mathbf{s}$ is vector through origin and $\mathbf{s}_m$, $\nabla \mathbf{f}(\mathbf{x}_i) \cdot \Delta \mathbf{x} + \mathbf{f}(\mathbf{x}_i)$ is the crossing between the tangent plane through $\mathbf{f}(\mathbf{x}_i)$ and the vector through $\mathbf{s}_m$. This can be used to find $\Delta \mathbf{x}$ and $\mathbf{x}_{i+1} = \mathbf{x}_i + \Delta \mathbf{x}$, and $\mathbf{f}(\mathbf{x}_{i+1})$ is the new function value. The algorithm stops when $\mathrm{norm}(\mathbf{f}(\mathbf{x}_{i+1}) - \beta \cdot \mathbf{s}) < \epsilon$, where ϵ is the stopping parameter. $\mathbf{f}(\mathbf{x})$ is the vector of detector signals if no fault would had occurred.

optical model, the cubic splined functions from [21] are used instead of the developed first principles models.

C. Fault Model

The purpose of the fault model and its related fault parameters are to use them for fault detection and classification. The structure of the combined optical and fault model, see (1), makes it possible to extract a pair of residuals which can be used for fault detection. In the following, one simple fault model is found.

In [21], two models are suggested, however, only the second seems to representing the fault. The model is illustrated in Fig. 6. This model uses two parameters to scale the four detector signals. One scales the focus detectors and another scales the radial detectors, see Fig. 6.

$$\mathbf{s}[n] = \begin{bmatrix} \beta_{\mathrm{f}}[n] \cdot \mathbf{I}_{2\times 2} & \mathbf{0} \\ \mathbf{0} & \beta_{\mathrm{r}}[n] \cdot \mathbf{I}_{2\times 2} \end{bmatrix} \cdot \mathbf{s}_{\mathrm{m}}[n] \qquad (24)$$

and the residuals based on $\beta_{\mathrm{f}}[n]$ and $\beta_{\mathrm{r}}[n]$ can be constructed as

$$\begin{bmatrix} \alpha_{\mathrm{f}}[n] \\ \alpha_{\mathrm{r}}[n] \end{bmatrix} = \begin{bmatrix} 1 - \beta_{\mathrm{f}}[n] \\ 1 - \beta_{\mathrm{r}}[n] \end{bmatrix}. \qquad (25)$$

These two residuals can be used for fault detection, but the major potential is in the classification process, since the two scalings, $\alpha_{\mathrm{f}}[n], \alpha_{\mathrm{r}}[n]$ would deviate for different faults. Two examples of faults can be seen in Section VII.

IV. DETERMINING THE INVERSE MAP OF THE OPTICAL MODEL

It is, as aforementioned, very interesting to estimate focus and radial distances, and the fault parameters based on the detector signals. A mapping from focus and radial distances, and fault parameters to the four detectors are given by the combined optical model and fault model. This mapping is nonlinear and not globally invertible. An estimate of focus and radial distances, and the fault parameters can be achieved by finding the inverse map of the combined optical and fault models. The method and algorithm are described in [21]. The method is based on the following facts. Due to the low-pass filtering nature of the OPU, the following limits are relevant: $|e_{\mathrm{f}}[n] - e_{\mathrm{f}}[n+1]| < \mu_{\mathrm{f}}$ and $|e_{\mathrm{r}}[n] - e_{\mathrm{r}}[n+1]| < \mu_{\mathrm{r}}$, where n is any sample, and μ_{f} and μ_{r} are the maximum changes in $e_{\mathrm{f}}[n]$ and $e_{\mathrm{r}}[n]$, see [21]. Based on the standard requirements to the CD player servos, see [19], the

maximum deviation of focus and radial positions from sample to sample can be calculated to: $\mu_{\mathrm{f}} = \mu_{\mathrm{r}} \approx 0.014 \ \mu\mathrm{m}$.

In the following, one approach is described to locally solve the inverse map and thereby find the right candidate point, and use this to compute the fault parameters and related residuals.

A. The Scaling Projection Method

The problem is to find the point, $\mathbf{f}(\mathbf{x}[n])$ in $\mathcal{D}$, where $\mathbf{x}[n] = [e_{\mathrm{f}}[n] \ e_{\mathrm{r}}[n]]^T$. By substituting the model output into (24) gives

$$\mathbf{f}(\mathbf{x}[n]) = \mathcal{B}[n] \cdot \mathbf{s}_{\mathrm{m}}[n] \qquad (26)$$

where

$$\mathcal{B}[n] = \begin{bmatrix} \beta_{\mathrm{f}}[n] \cdot \mathbf{I}_{2\times 2} & \mathbf{0} \\ \mathbf{0} & \beta_{\mathrm{r}}[n] \cdot \mathbf{I}_{2\times 2} \end{bmatrix} \in \mathcal{R}^{4\times 4}. \qquad (27)$$

Unfortunately, due to the shape of $\mathcal{D}$ in $\mathcal{R}^4$, there is a number of points in $\mathcal{D}$ for which (26) will be true. However, due to the system's low-pass-filtering effect $\mathbf{x}[n]$ is expected to be near $\mathbf{x}[n-1]$. This means that if an iterative algorithm is used to find $\tilde{\mathbf{x}}[n], \tilde{\mathbf{x}}[n-1]$ can be used as a starting value of the algorithm. Experiments have shown that it improves results if the dynamically estimated distances are used as this start value. For example, $\mathbf{x}_0[n] = [\hat{e}_{\mathrm{f}}[n] \ \hat{e}_{\mathrm{r}}[n]]^T$. The low-pass filtering properties of the estimator reduce noise in the estimates used as initial value of the iterations. The result is that the algorithm requires in general fewer iterations before the estimation error was as low as required.

In the following an algorithm, based on Newton's method, is described. One iteration of the algorithm used on the problem from Fig. 3, which is a problem in $\mathcal{R}^2$, is illustrated in Fig. 7.

First the crossing point between the tangent plane of $\mathcal{D}$ at the initial point and line which goes through origin and the measurement is found. The tangent plane of $\mathcal{D}$ at the initial point is

$$y_i[n] = \mathbf{f}(\mathbf{x}_i[n]) + \nabla \mathbf{f}(\mathbf{x}_i[n]) \cdot \Delta \mathbf{x}[n]$$
$$i = 0 \quad \text{and} \quad \mathbf{x}[n] \in \mathcal{R}^2. \qquad (28)$$

The tangent plane is a first-order approximation of $\mathcal{D}$ at the initial point. The intersection between this tangent plane and $\mathcal{B}[n] \cdot \mathbf{s}_{\mathrm{m}}[n]$ is consequently an approximation of (26). Finding the intersection, gives a $\Delta\mathbf{x}[n]$ and at a given iteration this gives an update in $\mathbf{x}[n]$ which is denoted $\Delta\mathbf{x}_{i+1}[n]$ This approximation can again be used as a starting point for a new tangent plane and a new intersection is computed. The intersection for iteration $i+1$ is

$$\mathbf{f}(\mathbf{x}_i[n]) + \nabla\mathbf{f}(\mathbf{x}_i[n]) \cdot \Delta\mathbf{x}_{i+1}[n] = \mathcal{B}[n] \cdot \mathbf{s}_{\mathrm{m}}[n] \Rightarrow \quad (29)$$

$$\nabla\mathbf{f}(\mathbf{x}_i[n]) \cdot \Delta\mathbf{x}_{i+1}[n] - \mathcal{B}[n] \cdot \mathbf{s}_{\mathrm{m}}[n] = -\mathbf{f}(\mathbf{x}_i[n]) \quad (30)$$

denote

$$\mathbf{S}_{\mathrm{m}}[n] = \begin{bmatrix} D_1[n] & 0 \\ D_2[n] & 0 \\ 0 & S_1[n] \\ 0 & S_2[n] \end{bmatrix} \in \mathcal{R}^{4 \times 2} \quad (31)$$

define the following two matrices

$$[\nabla\mathbf{f}(\mathbf{x}_i[n]) \quad -\mathbf{S}_{\mathrm{m}}[n]] \in \mathcal{R}^{4 \times 4}, \quad (32)$$

$$\begin{bmatrix} \Delta\mathbf{x}_{i+1}[n] \\ \beta_{\mathrm{f},i}[n] \\ \beta_{\mathrm{r},i}[n] \end{bmatrix} \in \mathcal{R}^{4 \times 1}. \quad (33)$$

This makes it possible to rewrite (30)

$$[\nabla\mathbf{f}(\mathbf{x}_i[n]) \quad -\mathbf{S}_{\mathrm{m}}[n]] \begin{bmatrix} \Delta\mathbf{x}_{i+1}[n] \\ \beta_{\mathrm{f},i}[n] \\ \beta_{\mathrm{r},i}[n] \end{bmatrix} = -\mathbf{f}(\mathbf{x}_i[n]) \quad (34)$$

The leftmost matrix in (34) is in practice invertible, since it has full rank, except from a few situations of low dimension. These situations can be avoided if the initial point is in the nominal range of operations.

$$\begin{bmatrix} \Delta\mathbf{x}_{i+1}[n] \\ \beta_{\mathrm{f},i}[n] \\ \beta_{\mathrm{r},i}[n] \end{bmatrix} = -([\nabla\mathbf{f}(\mathbf{x}_i[n]) \quad -\mathbf{S}_{\mathrm{m}}[n]])^{-1}$$
$$\cdot \mathbf{f}(\mathbf{x}_i[n]). \quad (35)$$

By using all four parameters to model the received detector signals, the suggested algorithm will converge as follows: $\lim_{i \to \infty} \mathbf{x}_i[n] = \mathbf{x}[n]$, $\lim_{i \to \infty} e_{\mathrm{f},i}[n] = e_{\mathrm{f}}[n]$, $\lim_{i \to \infty} e_{\mathrm{r},i}[n] = e_{\mathrm{r}}[n]$. This means it is relevant to use the L_2 norm of the difference between estimated and measured detector signals, as the stop criteria for the algorithm

$$\|(\mathbf{f}(\mathbf{x}_i[n]) - \mathcal{B}_i[n] \cdot \mathbf{s}_{\mathrm{m}}[n])\|_2 < \epsilon. \quad (36)$$

If the stop criteria is fulfilled, the algorithm stops, and if not fulfilled, the computed distances are used as initial guesses for the next iteration. When the algorithm is stopped the residuals are computed.

1) The Algorithm: Since the algorithm estimates $\mathbf{x}[n]$, the estimate is denoted $\tilde{\mathbf{x}}[n]$.

1) Find the gradient, $\nabla\mathbf{f}(\tilde{\mathbf{x}}_i[n])$ in the point $(\tilde{\mathbf{x}}_i[n], \mathbf{f}(\tilde{\mathbf{x}}_i[n]))$.
2) Compute

$$\begin{bmatrix} \tilde{\mathbf{x}}_{i+1}[n] \\ \beta_{\mathrm{f},i}[n] \\ \beta_{\mathrm{r},i}[n] \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{x}}_i[n] \\ 0 \\ 0 \end{bmatrix}$$
$$- ([\nabla\mathbf{f}(\tilde{\mathbf{x}}_i[n]) \quad -\mathcal{S}_{\mathrm{m}}[n]])^{-1} \cdot \mathbf{f}(\tilde{\mathbf{x}}_i[n]).$$

3) Compute $\gamma[n] = \|\mathcal{B}_i[n]\mathbf{s}_{\mathrm{m}}[n] - \mathbf{f}(\mathbf{x}_{i+1}[n])\|_2$

4) Jump to step 1 if $\epsilon < \gamma[n]$, else stop.
5) Compute the residuals $\begin{bmatrix} \alpha_{\mathrm{f}}[n] \\ \alpha_{\mathrm{r}}[n] \end{bmatrix} = \begin{bmatrix} 1 - \beta_{\mathrm{f},i}[n] \\ 1 - \beta_{\mathrm{r},i}[n] \end{bmatrix}$.

Regarding the convergence properties of the algorithm: The optical model is approximated by four products of second and third order polynomials. It is a known fact that the Newton-Raphson's method converges when solving scalar second and third order equations. However, since the model is not scalar, it is not known if convergence can be guaranteed, but convergence is assumed due to Newton-Raphson performance on scalar polynomials. In practice the algorithm has shown very good convergence, when tested. It has in every case converged within three iterations, with $\epsilon = 10^{-12}$.

B. Simulation of the Inverse Map Solution

In this section the ability of the algorithm to find the described inverse map is tested by a number of different simulations. The input signal to this simulation is chosen as two sine signals with a small difference in the frequency so that the two input signals are not fully correlated, see [21]. The frequency and amplitude of the these sine signals are chosen in a way such that the maximum difference value between two samples is at least: 0.014μm, see the beginning of Section IV. In the simulation a sample frequency at 35 kHz has been used. Starting with simulating the output of the optical model without any faults, see Fig. 8. The simulated faults are constructed based on the model in (24)–(25), where $\alpha_f[n]$ and $\alpha_r[n]$ model the scratches on the surface, and the signals are illustrated in Fig. 9. In this simulation $\alpha_f[n] = \alpha_r[n]$. Using the fault model and the fault signal in Fig. 9, the simulation series of samples with surface faults are computed and illustrated in Fig. 10. Before using the algorithm on these test data, independent random noise is added to the four detector signals, to simulate nondeterministic fault parts and noise. The next step in the simulation is to apply the signal from Fig. 10 to the algorithm. In this simulation the detector signals shown in Fig. 10 are fed to the algorithm. The results are shown in Fig. 11, which show $\mathbf{f}(\tilde{\mathbf{x}}[n])$. From these plots, it is clear that the scaling projection method is suited for finding this inverse map with the given fault model.

V. FAULT DETECTION

A commonly used strategy in handling surface faults like scratches is to use a fault tolerant control strategy, where the controller is accommodated to handle the fault when it is detected. Practical experience has shown that it is of great importance to make the fault detection as correct as possible.

The residuals based on the fault model can now be used for fault detection and in principle for fault classification. To illustrate the use of these residuals for possible fault detection, a threshold detection method is used. The threshold detection method has detected a fault at sample n if $f_{\mathrm{d}}[n] = 1$, and not detected a fault if $f_{\mathrm{d}}[n] = 0$. $f_{\mathrm{d}}[n]$ is defined in (37).

$$f_{\mathrm{d}}[n] = \max\{f_{\mathrm{d,f}}[n], f_{\mathrm{d,r}}[n]\} \quad (37)$$

where

$$f_{\mathrm{d},i}[n] = \begin{cases} 1 & \text{if } r_i[n] > \lambda_i \\ 0 & \text{if } r_i[n] < \lambda_i \end{cases} \quad (38)$$
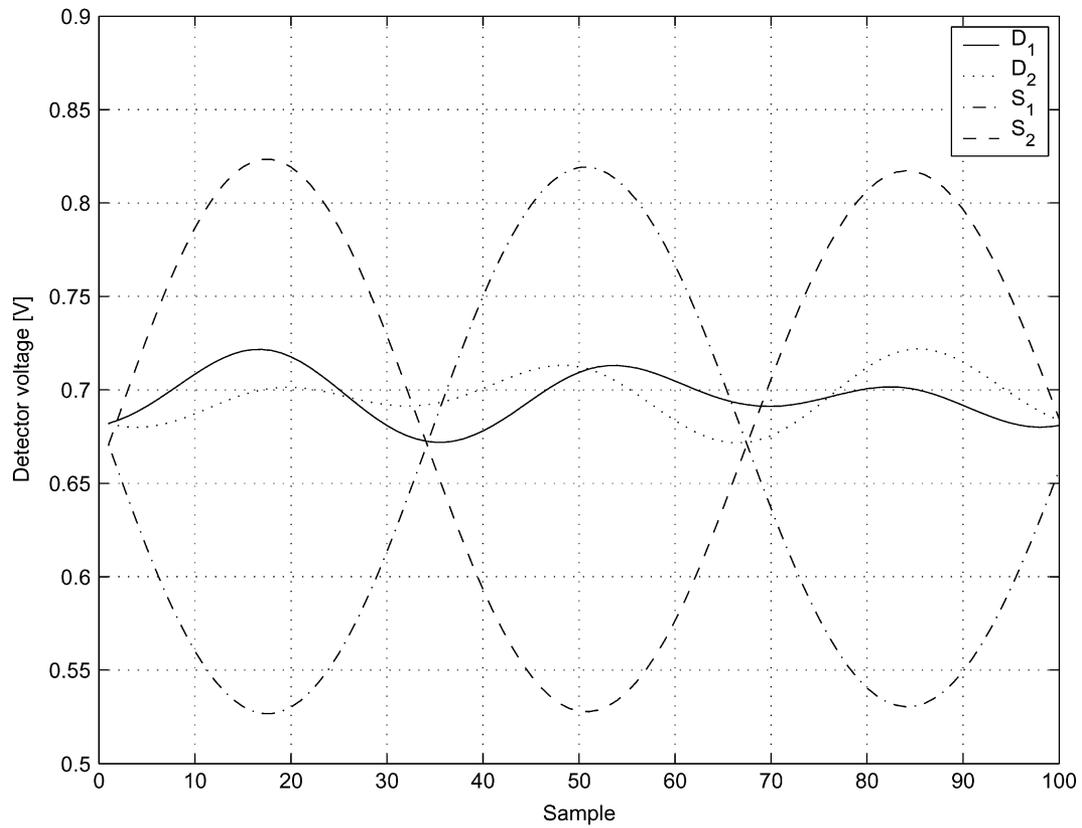
$$i \in \{\mathrm{f,r}\} \quad (39)$$

Fig. 8.   Simulated optical detector signals without faults used for simulation of the inverse map solver.
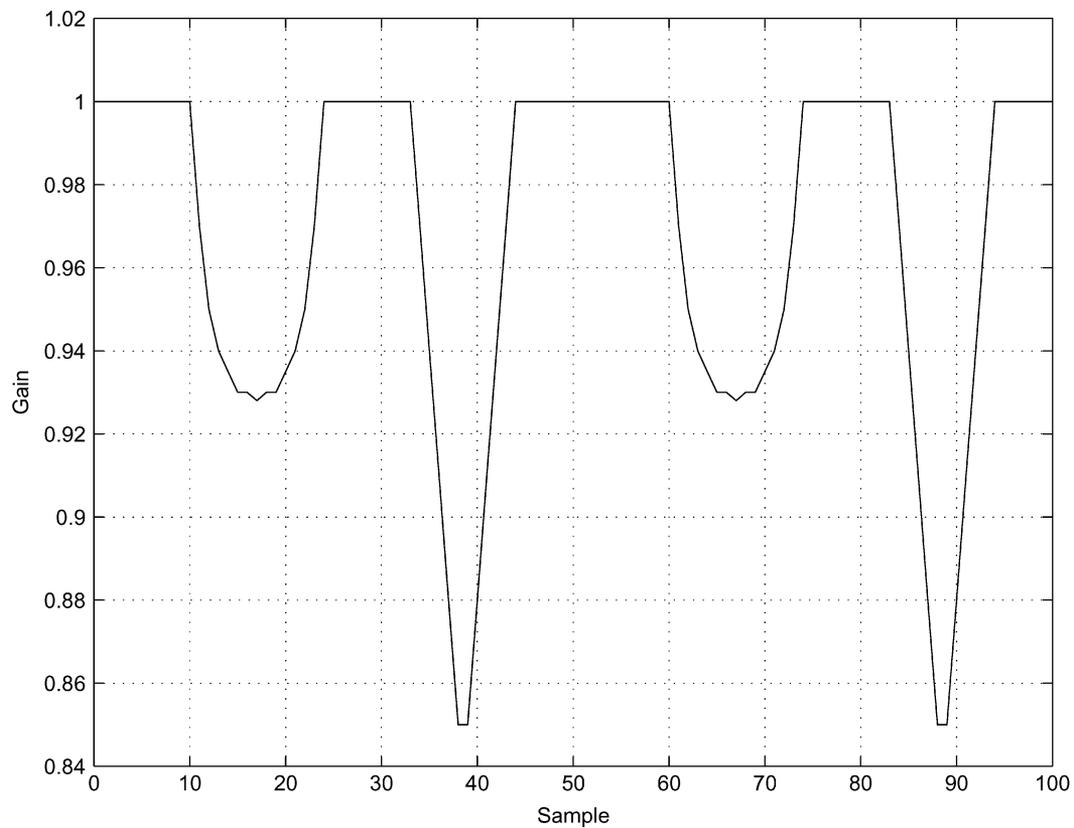

Fig. 9.   The $\alpha_f[n]$ and $\alpha_r[n]$ series, which are used as the surface faults in this simulation.
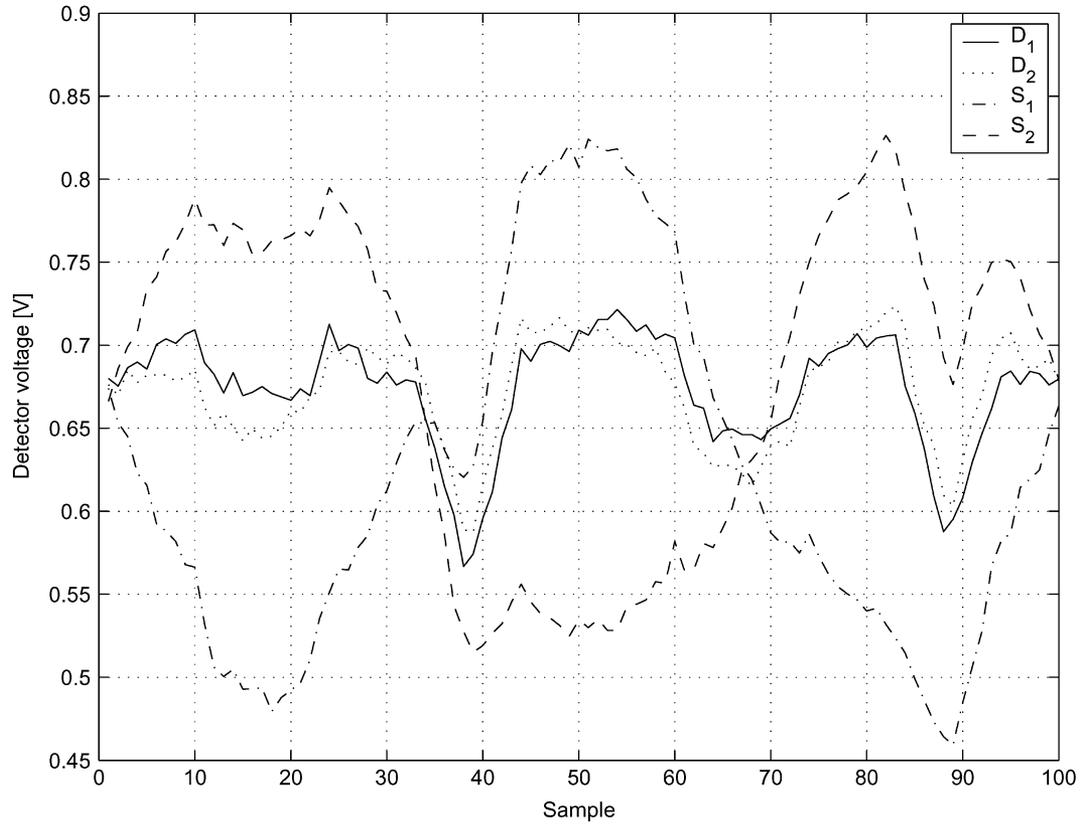
Fig. 10.   Simulation of the four detector signals $D_1[n]$, $D_2[n]$, $S_1[n]$ and $S_2[n]$ with surface faults and significant noise contribution.
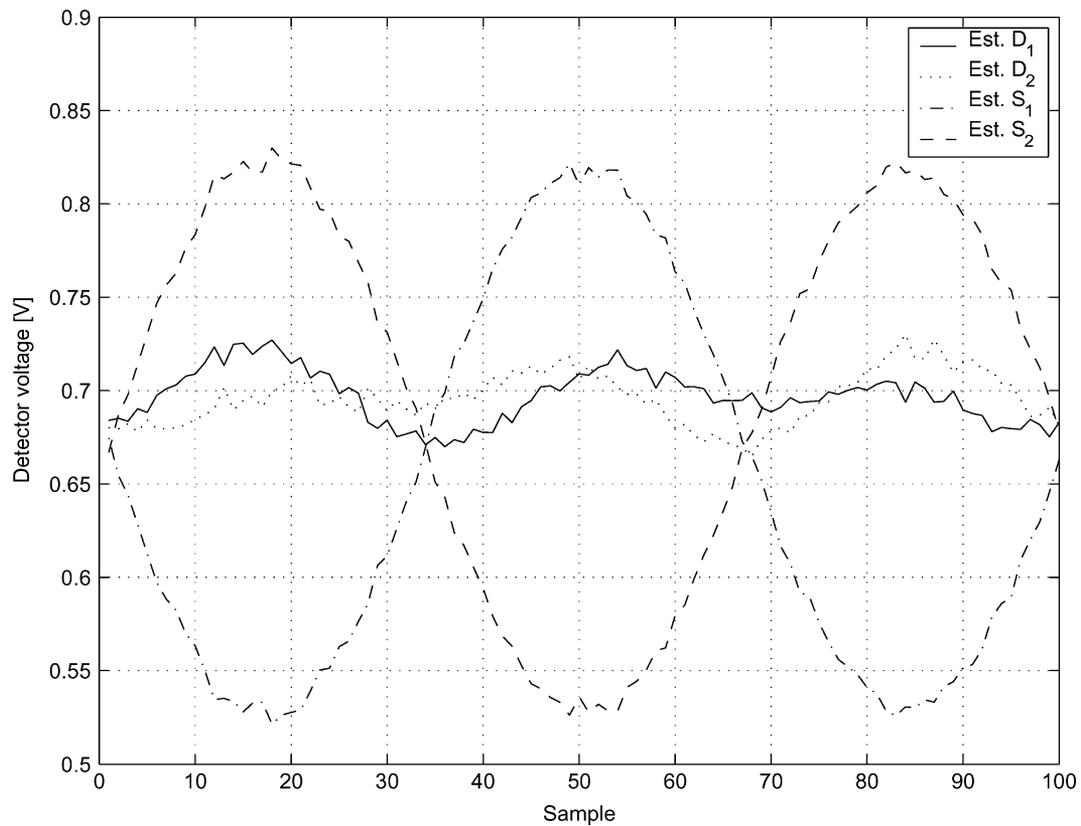

Fig. 11.   Illustration of the estimation of the four detector signals $D_1[n]$, $D_2[n]$, $S_1[n]$ and $S_2[n]$ by the use of the scaling projection method.
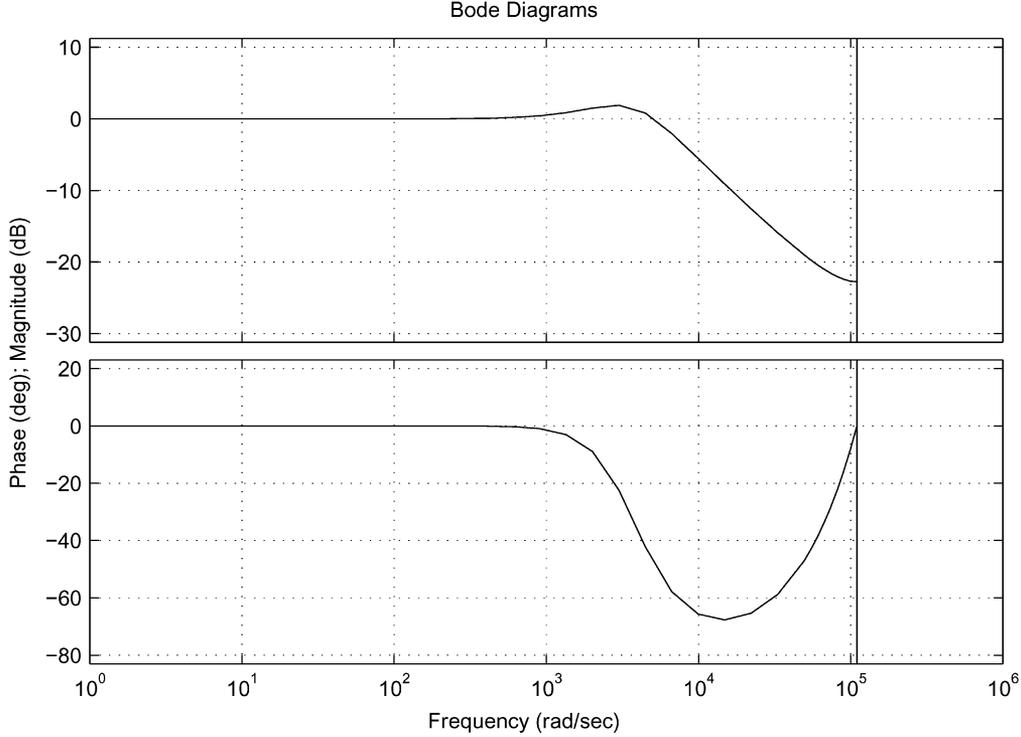
Fig. 12.   Bode plot of the Kalman estimator's transfer function from $\bar{e}_{\mathrm{f}}[n]$ to $\hat{e}_{\mathrm{f}}[n]$.

where $r[n]$ is the residual tested. In the tests, two set of residuals are tested $\{\alpha_{\mathrm{f}}[n], \alpha_{\mathrm{r}}[n]\}$ and $\{S_{\mathrm{f}}[n], S_{\mathrm{r}}[n]\}$. λ_i are the thresholds. $\{S_{\mathrm{f}}[n], S_{\mathrm{r}}[n]\}$ are the normally used residuals for surface fault detection—they are defined as follows: $S_{\mathrm{f}}[n] = D_1[n] + D_2[n]$ and $S_{\mathrm{r}}[n] = S_1[n] + S_2[n]$. This pair of residuals is also a representation of the light reflection ratio, but without taking cross-couplings and nonlinearities into account. This detection method is compared to a threshold method based on the normally used sum signals. These results are in the experiment section, see Section VII. The thresholds for the given signals are found such that they detect as much of the fault as possible and without making any false detections on the test data. These thresholds are found by a trial and error method.

VI. KALMAN ESTIMATOR

The static estimates of focus and radial distances also contain stochastic noise components. In order to make a more correct estimate of focus and radial distances subject to this noise a Kalman estimator is introduced.

Before designing a Kalman estimator the dynamical model of the electromagnetic system is discretised and expanded to include state and measurement noises as well. This gives the following model

$$\boldsymbol{\eta}[n+1] = \boldsymbol{\Phi} \cdot \boldsymbol{\eta}[n] + \boldsymbol{\Gamma} \cdot \mathbf{u}[n] + \boldsymbol{\Psi} \cdot \begin{bmatrix} w_1[n] \\ w_2[n] \end{bmatrix} \quad (40)$$

$$\begin{bmatrix} e_{\mathrm{f}}[n] \\ e_{\mathrm{r}}[n] \end{bmatrix} = \mathbf{C}_{\mathrm{CD}} \cdot \boldsymbol{\eta}[n] + \begin{bmatrix} v_1[n] \\ v_2[n] \end{bmatrix} \quad (41)$$

where $\boldsymbol{\Phi}, \boldsymbol{\Gamma}$ are the discrete time version of $\mathbf{A}_{\mathrm{CD}}, \mathbf{B}_{\mathrm{CD}}$. $w_1[n], w_2[n], v_1[n], v_2[n]$ are independent noises. In the design it is assumed that the $w_1[n], w_2[n]$ can be viewed as noises added to the control signal. The variance of $\mathbf{w}[n] = \begin{bmatrix} w_1[n] \\ w_2[n] \end{bmatrix}$ is named $\mathbf{Q}$, the variance of $\mathbf{v}[n] = \begin{bmatrix} v_1[n] \\ v_2[n] \end{bmatrix}$ is named $\mathbf{R}$, and the covariance is assumed to be zero.

These variance values are unknown and are consequently used as tuning parameters in the design process of the Kalman estimator. Due to the fact that the CD player will run for a very long time, it was chosen to design a steady state Kalman estimator, see [25]. That is the steady state Kalman gain, $\mathbf{L}_\infty$, and steady-state innovation gain, $\mathbf{M}_\infty$, are to be found. The estimator equations are seen in (42)–(45)

$$\hat{\eta}[n+1] = \boldsymbol{\Phi} \cdot \hat{\eta}[n] + \boldsymbol{\Gamma} \cdot \mathbf{u}[n] \quad (42)$$

$$+ \mathbf{L}_\infty \cdot \left(\begin{bmatrix} \tilde{e}_{\mathrm{f}}[n] \\ \tilde{e}_{\mathrm{r}}[n] \end{bmatrix} - \mathbf{C}_{\mathrm{CD}} \cdot \hat{\eta}[n] \right) \quad (43)$$

$$\begin{bmatrix} \hat{e}_{\mathrm{f}}[n] \\ \hat{e}_{\mathrm{r}}[n] \end{bmatrix} = \mathbf{C}_{\mathrm{CD}} \cdot (\mathbf{I} - \mathbf{M}_\infty \cdot \mathbf{C}_{\mathrm{CD}}) \cdot \hat{\eta}[n] \quad (44)$$

$$+ \mathbf{C}_{\mathrm{CD}} \cdot \mathbf{M}_\infty \cdot \begin{bmatrix} \tilde{e}_{\mathrm{f}}[n] \\ \tilde{e}_{\mathrm{r}}[n] \end{bmatrix}. \quad (45)$$

The two gains are calculated by use of the Matlab function *Kalman*. This function has the variance matrices as inputs, and they are found during an iterative design process of the Kalman estimator, such that the Kalman estimator does not remove the low frequently disturbances. The output of this Kalman estimator is illustrated for an experimental example in Fig. 16. In
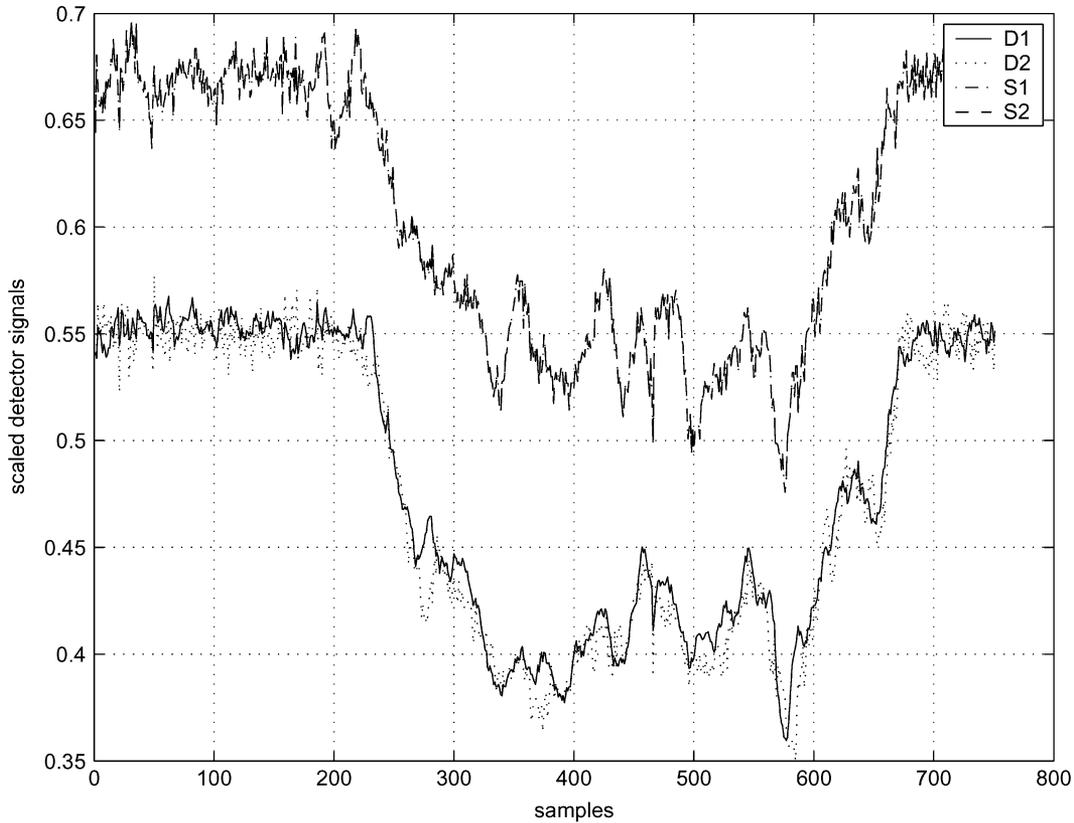
Fig. 13. Measured detector signals $D_1[n]$, $D_2[n]$, $S_1[n]$, and $S_2[n]$ while passing the scratch.

Fig. 12 a Bode plot of the Kalman estimators transfer function from $\tilde{e}_f[n]$ to $\hat{e}_f[n]$ is shown. Experimental work indicates that disturbances are not significantly present above 15 krad/s. This is reflected by the high frequent roll off in the Kalman filter Bode plot in Fig. 12. The same experimental work indicates that surface faults have frequency content in the range from 3 krad/s. The gain of the observer is only high in the frequency range for which the disturbances are present. However, a part of this range is also in the frequency range of the surface faults. This means that the estimator contains the disturbances and removes a large part of the fault signal from the estimate.

VII. EXPERIMENTAL DATA

The experimental setup consists of a CD player, with a three beam single Foucault detector principle, a PC with an I/O-card, and some hardware in order to connect the CD player with the I/O-card. Due to the limited computational power of the CPU in the PC the sample frequency is chosen to 35 kHz. This is lower than the normal CD servo sample frequency (44 kHz). The four detector signals and the two control signals are sampled. By using the built-in controller of the CD player, a number of CDs with certain faults are sampled in a normal operation. Normal operation means that the faults are not severe enough to force the CD player in a state where it cannot play a certain disc, but on the other hand the faults are challenging for the built-in controllers and simple fault handling algorithm.

The experimental work has mainly been focused on real scratches and fingerprints, since artificial faults tend to be nice and regular, and they are not really challenging. For example

some artificial faults are equivalent with a step disturbance in the focus loop, which the controller should reject without any trouble.

The method has been used on a number of different faults, and they all show similar results to the two presented in this paper. The method is applied to a scratch. For the scratch five plots are shown, illustrating the sampled signals, see Fig. 13. The estimated distances $(\hat{e}_f[n], \hat{e}_r[n])$ can be seen in Fig. 16. The fault detection based on the focus residuals can be seen in Fig. 14. The fault detection based on the radial residuals can be seen in Fig. 15.

VIII. DISCUSSION ON THE ESTIMATED SIGNALS

In the following, Figs. 13–17 will be commented. Starting with Fig. 13 which illustrates the scaled detector signals, since it is easy from this figure to do a visual detection of the scratch. The real scratch is the part of the signals where the values are decreased. It lasts approximately from sample 230 to sample 670. The ideal fault detection algorithm will end up with detections of beginnings and ends, which, however, is difficult to obtain in practice. Figs. 14 and 15 illustrate the beginning and end of the scratch seen in Fig. 13, since practical experience has shown that the detection of the beginning and end of the surface fault is important.

From Figs. 14 and 15 it can be seen that $S_f[n]$, $S_r[n]$, $\alpha_r[n]$ and $\alpha_r[n]$ are good fault parameters/residuals for detection of the fault, since the size of the variations in the residuals during the scratch are significantly larger than the variations due to noise outside the scratch. Based on these residuals a threshold
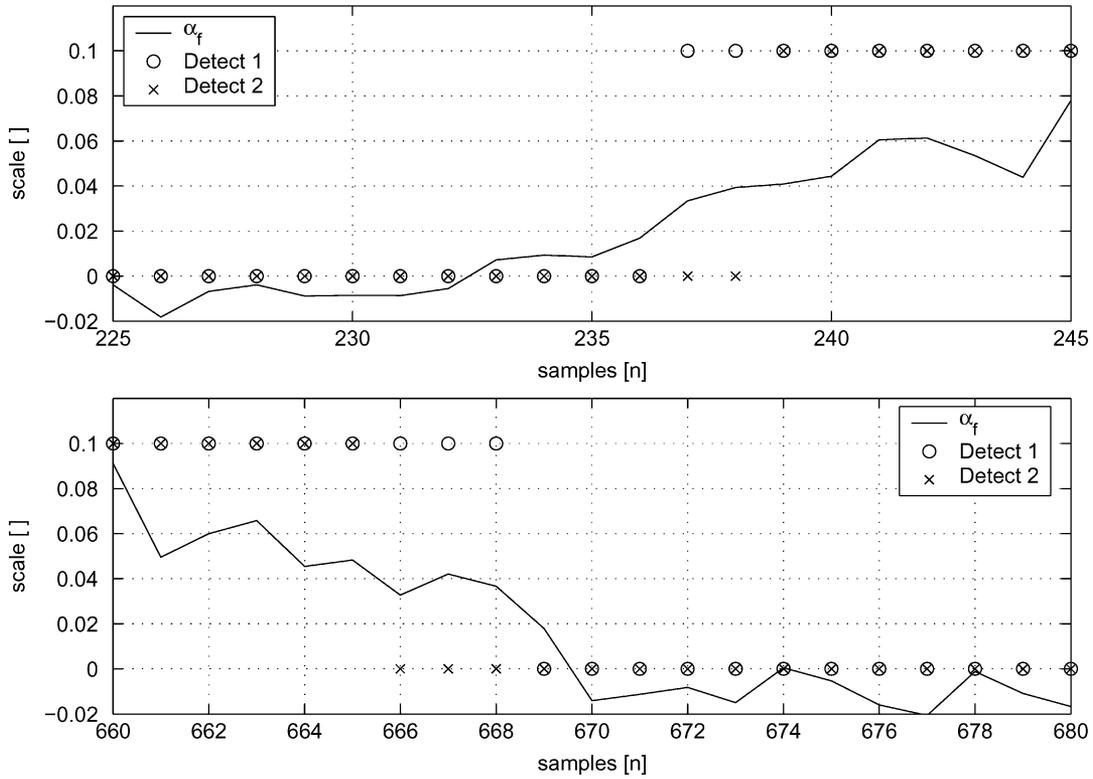
Fig. 14. $\alpha_f[n]$ plotted together with detection signals based on both $\alpha_f[n]$ (Detect 1) and $S_f[n]$ (Detect 2). The upper plot shows the beginning detection of the scratch, and the lower plot shows the end detection.
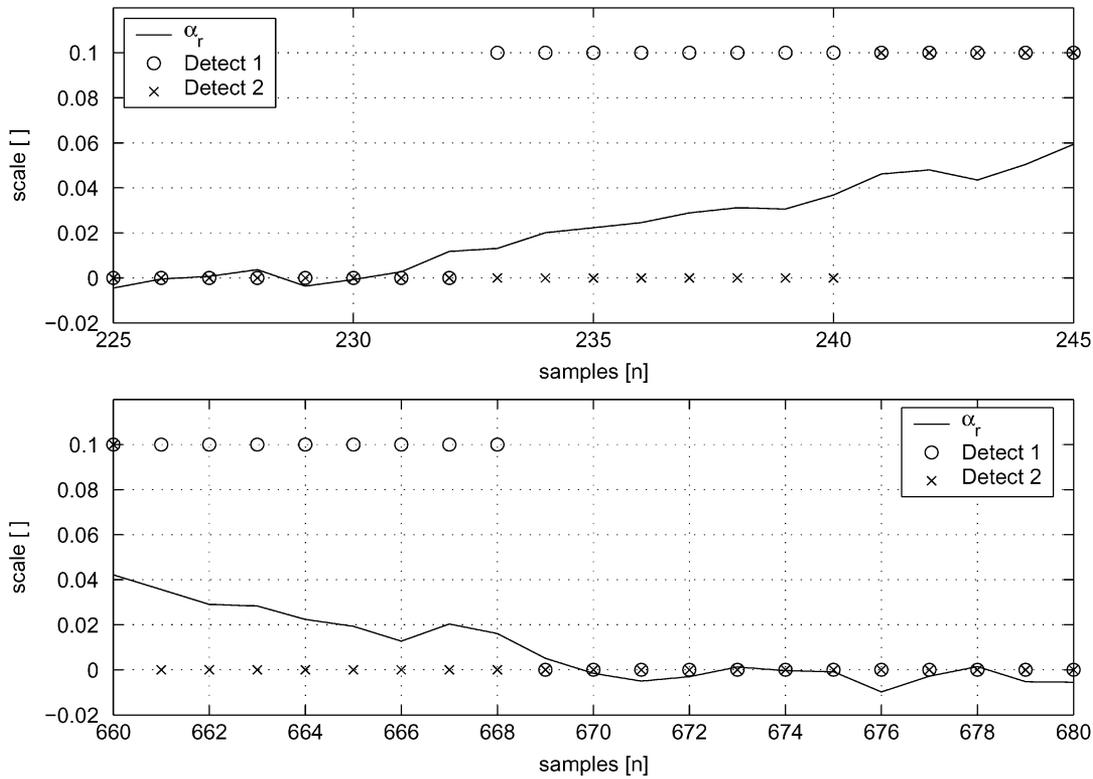

Fig. 15. $\alpha_r[n]$ plotted together with detection signals based on both $\alpha_r[n]$ (Detect 1) and $S_r[n]$ (Detect 2). The upper plot shows the beginning detection of the scratch, and the lower plot shows the end detection.

detection is implemented, by using a threshold which gives the earliest detection of the scratch without giving false detections, (the thresholds were found by lowering these thresholds until false detections were present). The detections based on the four
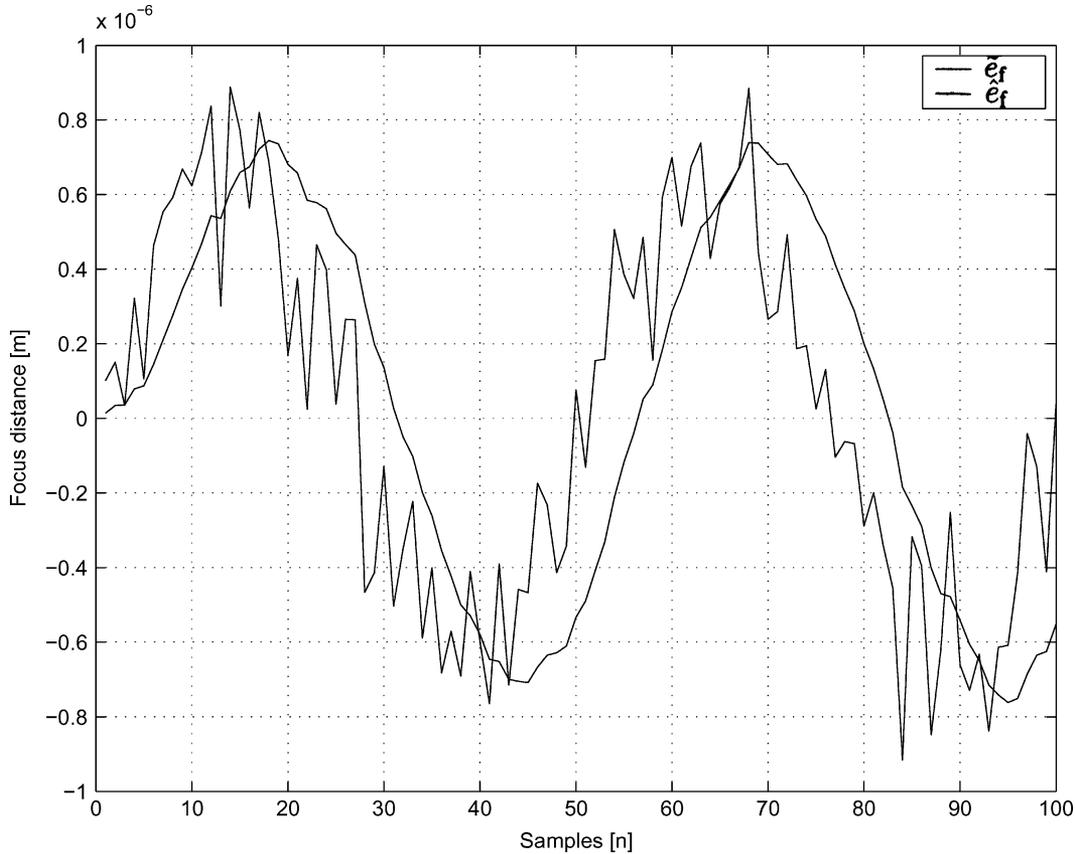
Fig. 16. The statically and dynamically estimated focus distance signals used in the simulation of the inverse mapping solver.

residuals are for: the αs: $n = [233–668]$, the Ss: $n = [239–665]$. In industrial applications often only one of the sum signals is used to perform the fault detection. In this example the worst case scenario would be if only the radial sum signal is used. If only the radial sum is used to detect the fault, it will result in a detection in the interval $S_r : n = [241–660]$. From these it can be seen that fault detection based on the new decoupled residuals $\alpha_f[n]$ and $\alpha_f[n]$ give a more clear detection than if the two other residuals were used, since the background noise level is lower. The results using the new residuals show an improvement by 6 samples in the beginning and 3 samples in the end. In [26] an extention of the threshold method used on the new residuals is presented, using this method results in an avarage improvement of the fault localization of 6–8 samples in both the begining and the end of the fault. This improvement can seem less significant compared to duration of the scratch which is 441 samples, but many practical experiments with fault tolerant control in CD players have shown that an improvement of a few samples of the fault detection means a lot in the achieved controller performance. Studies have indicated that a delayed detection of the beginning of the surface fault will result in a control error which depends quadratically on the number of samples of the delay.

One should also note that the new method for detection is very close to the fault localization by the visual inspection. It achieves an error of 3 samples in the beginning and only 2 samples in the end. It is a clear improvement if compared with the fault detection based on the normal residuals which has an error of 9 samples in the beginning and 5 in the end.

In addition $\boldsymbol{\alpha}_f[n]$ and $\boldsymbol{\alpha}_r[n]$ are "by construction" decoupled from the respective radial and focus distances, (see [21]), which the sum signals (the normal used residuals) are not. In the example with sum signals, the thresholds were not chosen such that the detection is robust toward the optical cross-couplings, imply that if the detection should be robust toward the cross-couplings the threshold would increase. This in turn would imply that the beginning of the fault would be detected later and the end earlier.

Fig. 16 shows $\hat{e}_f[n]$ and $\tilde{e}_f[n]$ in the case of the signal used to test the inverse map solver. From this figure it can be seen that the Kalman estimator removes the noise from the static estimate of focus distance but introduces a phase shift. The reason is that the original test signal is outside the normal frequency region of disturbances, and as a consequence the Kalman estimator is designed with a phase shift at the frequency, see Fig. 12.

In order to evaluate these estimates on actual experimental data, another illustration is needed. Fig. 17 illustrates a zoom on $\hat{e}_f$ and $\tilde{e}_f$ where a part of the scratch is contained. From this illustration it can be seen that a large part of the scratch is removed from $\tilde{e}_f$ to $\hat{e}_f$ due to the Kalman filter. A part of the signal due to the scratch is still contained in the signal. Since it is in the frequency range of the disturbances, it has to be handled by the use of combined time-frequency filtering. Inspecting the radial estimates gives the same results. The inspection of the estimates leads to the conclusion that $\hat{e}_f[n]$ is a better estimate of the real distance than $\tilde{e}_f[n]$. However, it is not possible to validate the estimation by comparing them with real distances, since it is not
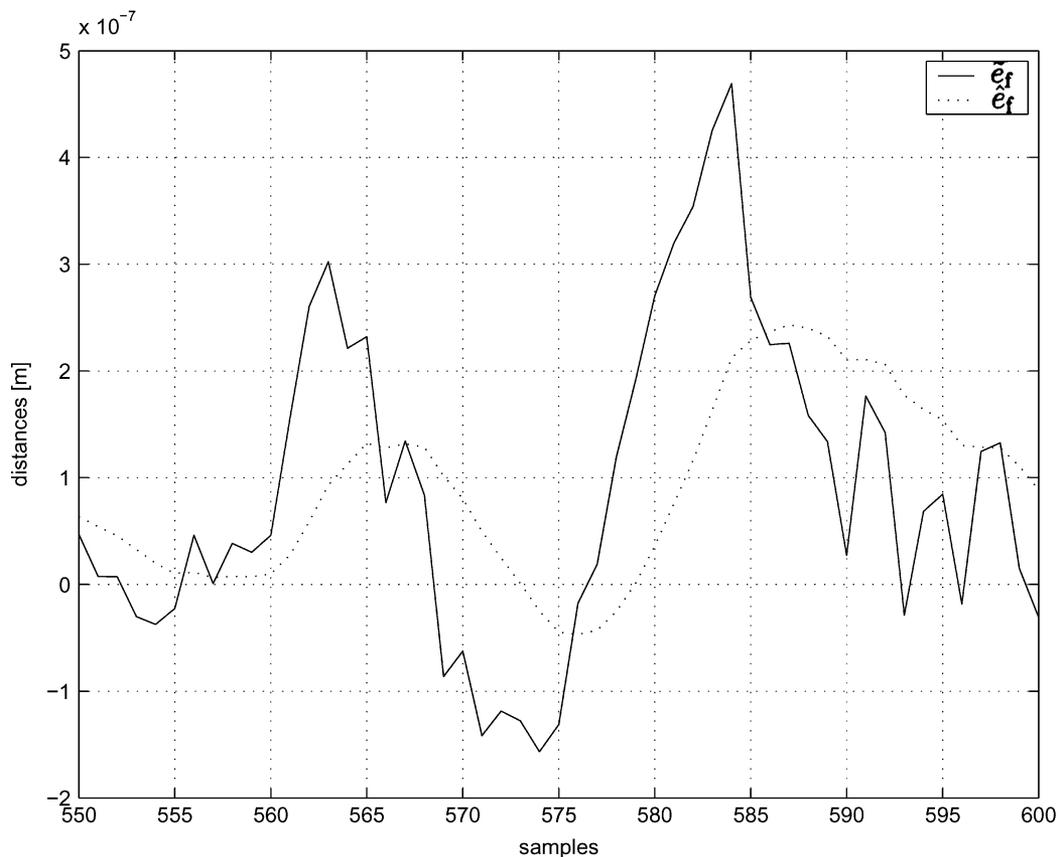
Fig. 17. A zoom on $\hat{e}_f$ and $\bar{e}_f$ where a part of the scratch is contained. Noticed that a large part of the scratch is removed from $\bar{e}_f$ to $\hat{e}_f$ due to the Kalman filter.

possible to measure those. Notice also the relatively large values of $\hat{e}_f[n]$ and $\hat{e}_r[n]$ in the beginning and especially the end of the scratch. These variations are due to heavy controller activations, indicating that the controller reacts on faulty sensor signals, and in case where the reaction is in the end of the fault it reacts on wrong state informations.

IX. CONCLUSION

This paper introduces a new method for estimating fault residuals, and focus and radial distances in a CD player. The method consists of four parts, the optical model, a solution to the inverse map of the optical model, a Kalman estimator for dynamically estimating focus and radial distances, and a fault residual extraction part. In order to estimate the new decoupled residuals and estimates of focus and radial distances, a model of the optical detector system is made, as well as a fault model. Together, these two models form a nonlinear and nonglobally invertible mapping from residuals, focus and radial distances to the four measured detector signals. A method to compute the inverse of the mapping is made based on Newton-Raphson's method. Finally, a Kalman estimator is designed to make a better estimate of focus and radial distances. The method is applied to experimental data, containing a scratch. By using threshold detection on the fault residuals the fault detection is improved for a given scratch. The suggested method detects the beginning of this scratch earlier than the normally used threshold method on the sum signals method. In addition it should be noticed that

the conventional sum signals method is not robust toward focus and radial cross-coupling. The method proposed in this paper reduces the cross-couplings between focus and radial loops. Focus and radial distances are estimated, and seems to be better estimates of the actual distances compared to the normally used indirect measures. The new estimate of the distances removes optical cross-couplings and a part of the frequency content of the faults. However, it is not possible in the experimental setup to measure the actual distances, and therefore not possible to validate estimated distances by comparing them to the real ones.

REFERENCES

[1] "Product specification: Digital servo processor DSIC2, TDA1301T," in *Philips Semiconductors*. Eindhoven, The Netherlands: Philips, Mar. 1994.

[2] P. Andersen, T. Pedersen, J. Stoustrup, and E. Vidal, "Method for improved reading of digital data disc," Int. patent WO 02/05 271 Al, 2001.

[3] E. Vidal, K. Hansen, R. Andersen, K. Poulsen, J. Stoustrup, P. Andersen, and T. Pedersen, "Linear quadratic control with fault detection in compact disk players," in *Proc. 2001 IEEE Int. Conf. Control Applications*, Mexico City, Mexico, 2001.

[4] M. Steinbuch, G. Schoostra, and O. Bosgra, "Robust control of a compact disc player," in *Proc. 31st IEEE Conf. Decision and Control*, Tucson, AZ, Dec. 1992.

[5] W. Draijer, M. Steinbuch, and O. Bosgra, "Adaptive control of the radial servo system of a compact disc player," *Automatica*, vol. 28, no. 3, pp. 455–462, May 1992.

[6] G. Dötch, T. Smakman, P. Van den Hof, and M. Steinbuch, "Adaptive repetitive control of a compact disc mechanism," in *Proc. 34th IEEE Conf. Decision and Control*, New Orleans, LA, 1995.

[7] G. Hearns and M. Grimble, "Limits of performance of an optical disk drive controller," in *Proc. 1999 American Control Conf.*, San Diego, CA, Jun. 1999.

[8] J. Li and T.-C. Tsao, "Rejection of repeatable and nonrepeatable disturbances for disk drive actuators," in *Proc. 1999 American Control Conf.*, San Diego, CA, Jun. 1999.

[9] J.-Y. Yen, C.-S. Lin, Y.-Y. Li, and C.-H. Chen, "Servo controller design for an optical disk drive using fuzzy control algorithm," in *IEEE Int. Conf. Fuzzy Systems*, San Diego, CA, Mar. 1992.

[10] L. Yao, A.-M. Wang, and Y.-F. Cheng, "Track seeking hybrid fuzzy controller for the compact disc player," in *The 10th IEEE Int. Conf. Fuzzy Systems*, Melbourne, Australia, Dec. 2001.

[11] S. Weerasooriya and D. Phan, "Discrete-time LQG/LTR design and modeling of a disk drive actuator tracking servo system," *IEEE Trans. Ind. Electron.*, vol. 42, no. 3, pp. 240–247, Jun. 1995.

[12] K. Fujiyama, M. Tomizuka, and R. Katayama, "Digital tracking controller design for cd player using disturbance observer," in *Int. Workshop on Advanced Motion Control*, Coimbra, Portugal, Jun. 1998.

[13] J. W. Heo and J. Chung, "Vibration and noise reduction of an optical disk drive by using a vibration absorber," *IEEE Trans. Consum. Electron.*, vol. 48, no. 4, pp. 874–878, Nov. 2002.

[14] P. Odgaard, J. Stoustrup, P. Andersen, and H. Mikkelsen, "Modeling of the optical detector system in a compact disc player," in *Proc. 2003 American Control Conf.*, Denver, CO, Jun. 2003.

[15] W. Bouwhuis, J. Braat, A. Huijser, J. Pasman, G. van Rosmalen, and K. S. Immink, *Principles of Optical Disc Systems*. New York: Adam Hilger, 1985.

[16] E. Vidal, J. Stoustrup, P. Andersen, T. Pedersen, and H. Mikkelsen, "Open and closed loop parametric system identification in compact disk players," in *ACC2001*, Arlington, VA, 2001.

[17] T.-J. Yen and Y.-C. Pan, "Modeling and identification of opto-mechanical coupling and backlash nonlinearity in optical disk drives," *IEEE Trans. Consum. Electron.*, vol. 46, no. 1, pp. 105–115, Feb. 2000.

[18] M. Dettori, "LMI techniques for control-with application to a compact disc player mechanism," Ph.D. dissertation, Technische Universiteit Delft, The Netherlands, 2001.

[19] S. G. Stan, *The CD-ROM Drive*: Kluwer Academic Publishers, 1998.

[20] E. Vidal, P. Andersen, J. Stoustrup, and T. Pedersen, "A study on the surface defects of a compact disk," in *Proceedings of the 2001 IEEE International Conference on Control Applications*, Mexico City, Mexico, 2001.

[21] P. Odgaard, J. Stoustrup, P. Andersen, and H. Mikkelsen, "Fault detection for compact disc players based on redundant and nonlinear sensors," in *Proc. American Control Conf. 2004*, Boston, MA, Jun.–Jul. 2004.

[22] ——, "Extracting focus and radial distances, fault features from CD player sensor signals by use of a kalman estimator," in *Proc. 42nd IEEE Conf. Decision and Control*, Maui, HI, Dec. 2003.

[23] M. Steinbuch, P. van Groos, G. Schootstra, and O. Bosgra, "Multivariable control of compact disc player using DSPs," in *Proc. American Control Conf.*, Baltimore, MD, Jun. 1994.

[24] *The Compact Disc Handbook*, 2nd ed., Oxford University Press, Oxford, U.K., 1992. K. Pohlmann.

[25] G. F. Franklin, J. D. Powell, and M. Workman, *Digital Control of Dynamic Systems*, 3rd ed. Reading, MA: Addison-Wesley, 1998.

[26] P. Odgaard and M. Wickerhauser, "Time localization of surface defects on optical discs," in *Proc. 2004 IEEE CCA/ISIC/CACSD*, Taipei, Taiwan, Sep. 2004.

Jakob Stoustrup (S'87–M'91–SM'99) was born in 1963 in Aalborg, Denmark. He received the M.Sc.E.E. degree from the Technical University of Denmark in 1987. He received the Ph.D. degree in applied mathematics in 1991.

He was a Teaching Assistant with the Technical University of Denmark, and a Visiting Researcher with Eindhoven University of Technology, The Netherlands, in 1988. In 1991, he was a Senior Researcher with the Danish Technical Research Council. During 1991–1995, he was an Assistant Professor and an Associate Professor during 1995–1996, both at the Department of Mathematics, Technical University of Denmark. He was a Visiting Professor at the University of Strathclyde, Glasgow, U.K., in 1996. Since 1997, he has been a (full) Professor with the Department of Control Engineering, Aalborg University, and since 2003, a Visiting Professor with the Institut Mittag-Leffler, Sweden. He is the author of more than 100 papers and four international patents. His research areas include robust control, fault diagnosis, and fault-tolerant control.

Dr. Stoustrup is a Member of the Governing Board and Council of the European Union Control Association (EUCA), a Member of IFAC Technical Committees: SAFEPROCESS and Robust Control. He is Past Chairman for the IEEE Joint Chapter: CSS+RAS; Member of The Danish Academy of Technical Sciences, 2005. He received the Statoil Prize.

Palle Andersen received a B.Sc. in electrical engineering from Danmarks Ingenirakademi, Aalborg, in 1974. From 1975 to 1978 he was with Nellemann A/S working with optimization of waste water plant operation using automation in a research project in association with Aalborg University and in 1977 he received the Industrial Ph.D. (erhvervsforsker) from Aalborg University. From 1978 to 1986 he was with the municipality in Randers working with automatic control of district heating, water supply, and production and distribution of power. From 1986 he has been with Aalborg University, from 1991 as a associate professor. He has been working with modeling and control in energy systems like power plant boilers and district heating systems and has received the IFAC Control Engineering Practice Prize Paper Award. Another research topic has been control electro-mechanical systems including CD players and he holds a patent on a method for improved reading of digital data discs.

Peter Fogh Odgaard (S'03–M'06) received the M.Sc.E.E. and Ph.D. degrees in electrical engineering with specialization in control engineering from Aalborg University, Denmark, in 2001 and 2004, respectively.

During his Ph.D. work on fault-tolerant control of CD players, he was a short-time guest researcher with Philips Research, Eindhoven, The Netherlands, in 2002, and a Guest Researcher with Washington University, St. Louis, MO, in 2003. He is currently a Research Assistant Professor with the Department of Control Engineering, Aalborg University, working with fault detection and performance monitoring of fossil-fired-power plants. His research interests include fault detection, isolation, and accommodation, performance monitoring and prediction, optimal and robust control, wavelets, control of optical disk drives, and control of power plants.

Henrik Fløe Mikkelsen was born in 1964 in Denmark, and recieved his B.Sc. EE. in 1989 from Aarhus High School of Engineering Research Engineer. He has been with Bang & Olufsen from 1990. He has from 1990 to 1996 been working on adaptive signal processing and control, pattern recognition, neural networks Technology Specialist in Control Engineering and optical storage. He has in 1996–2004 been working on optimizing CD, DVD players tracking control system. From 2004 has been Head of Audio Technology & Innovation; in charge of research, technologies and innovations to be used in audio products, Radio, CD and DVD players and also for surround sounds systems.