A Dynamical Systems Approach to Failure Prognosis

David Chelidze
Department of Mechanical Engineering &
Applied Mechanics,
University of Rhode Island,
Kingston, RI 02881
e-mail: chelidze@egr.uri.edu;
http://www.mce.uri.edu/chelidze/

Joseph P. Cusumano
Department of Engineering Science &
Mechanics,
Pennsylvania State University,
University Park, PA 16802
e-mail: jpc@crash.esm.psu.edu;
http://www.esm.psu.edu/nld/

In this paper, a previously published damage tracking method is extended to provide failure prognosis, and applied experimentally to an electromechanical system with a failing supply battery. The method is based on a dynamical systems approach to the problem of damage evolution. In this approach, damage processes are viewed as occurring in a hierarchical dynamical system consisting of a "fast," directly observable subsystem coupled to a "slow," hidden subsystem describing damage evolution. Damage tracking is achieved using a two-time-scale modeling strategy based on phase space reconstruction. Using the reconstructed phase space of the reference (undamaged) system, short-time predictive models are constructed. Fast-time data from later stages of damage evolution of a given system are collected and used to estimate a tracking function by calculating the short time reference model prediction error. In this paper, the tracking metric based on these estimates is used as an input to a nonlinear recursive filter, the output of which provides continuous refined estimates of the current damage (or, equivalently, health) state. Estimates of remaining useful life (or, equivalently, time to failure) are obtained recursively using the current damage state estimates under the assumption of a particular damage evolution model. The method is experimentally demonstrated using an electromechanical system, in which mechanical vibrations of a cantilever beam are dynamically coupled to electrical oscillations in an electromagnet circuit. Discharge of a battery powering the electromagnet (the "damage" process in this case) is tracked using strain gauge measurements from the beam. The method is shown to accurately estimate both the battery state and the time to failure throughout virtually the entire experiment. [DOI: 10.1115/1.1640638]

1 Introduction

Most previous work in the field of machinery condition monitoring has aimed at developing robust discriminators of *impending* failures. Current research efforts, however, aim to move past such diagnostic methods, which are primarily suitable for alarm-based condition monitoring systems, to the actual tracking of incipient damage as it develops. Such "gray scale" damage state trackers are required, furthermore, for the development of true prognostic algorithms capable of giving continuously updated estimates of remaining life, well in advance of actual failures.

In this paper we formulate and experimentally implement a new method for failure prognosis that addresses this need. The work presented here is an extension of previous efforts [1,2] to develop a method for tracking hidden damage processes. The diagnostic and prognostic methodology employed is based on a dynamical systems perspective on evolving damage [3]. From this point of view, a slowly evolving damage process is coupled to, and causes nonstationarity in, a directly observable fast-time dynamical system. Because of this coupling, appropriate determination of subtle distortions in the state-space dynamics of the fast time system can be used to track the developing damage and predict its future evolution.

Here, the utility of the method is validated experimentally by the tracking and prediction of a battery discharge process in an electro-mechanical system. The application of these methods to real damage processes in structures (crack initiation and progression in a structure) is being explored in current research. Preliminary results for material damage were presented in [2] and final results will be the subject of forthcoming papers.

In the next section, a current literature review is used to frame our approach to damage diagnosis and prognosis. A brief description of our previously presented damage tracking procedure is given in Section 3. In Section 4, we develop procedures for improved damage diagnosis and prognosis using recursive filters. The application of the proposed techniques to the experimental electromechanical system is described in Section 5. Finally, in Section 6 we conclude with a discussion of our results.

2 Background

Review articles that discuss the state-of-the-art of damage identification and systems health monitoring are presented by Doebling et al. [4,5], and more recently by Zou et al. [6]. In the following paragraphs we give a brief overview of some of the main developments in damage diagnosis and prognosis that provide a context and background for the work presented in this paper.

Much of the research on damage detection and identification has concentrated on the development of heuristic methods based on time or frequency domain signal processing techniques [7–10]. More recent research focuses on heuristic methods that can utilize both time and frequency domain information [11–13]. In both cases, these are mainly failure detection methods, i.e. the main emphasis is on the development of a feature vector that will indicate when the system parameters have reached some preset failure values. These methods do not continuously estimate the damage (or, equivalently, health) state, or provide a functional relationship between the feature vector and the damage state—a necessary requirement for failure prognosis. The main advantage of these methods is that they are easy to implement and sometimes work very well. However, there is no theoretical basis to determine *a priori* if a given method is going to work well for a particular system without prior experimentation or knowledge.

Model based methods, in contrast, overcome some of the limitations of heuristic methods at the expense of more complex development and higher implementation costs [14–16]. They are general in the sense that if some properties of the system or damage physics change, models can be readjusted to accommodate the change. Their main advantage is that knowing a model structure gives one the ability to tie the changes in feature vectors to the

model parameter changes [17]. However, in many cases one does not know the appropriate model, and, since most damage processes are nonlinear, new model development costs are high. This problem is overcome by the use of autoregressive models [18], neural networks [19–21], genetic algorithms [22], and other data-based modelling techniques in place of physics-based models. However, the ability to directly tie damage evolution to changes in a system's physical parameters is generally lost with data-based models. It should be mentioned that most of the model based schemes, just like most heuristic methods, have up to now been used primarily to provide damage *detection* and, possibly, classification.

The failure prognosis problem is still in its developmental stages. Currently available failure prediction algorithms can be divided into methods based on deterministic [23] and probabilistic or stochastic [24,25] models for fault propagation. Currently, these methods are application dependent and their success is contingent on a successful damage *state estimation* by some means. Damage state estimation (or *damage tracking*) refers to a process that goes beyond mere identification and classification of damage to yield a continuous, updateable estimate of the current state of damage in a given system or component, sometimes referred to as "gray scale" damage monitoring.

This paper is based on the assertion that for true prognostic ability we need: (1) an empirical or physics-based damage evolution model; (2) a failure surface (or levels) for the relevant damage variables that operationally defines the failed condition; and (3) a technique that can continuously estimate the current state of these variables. It is preferable that such a damage tracking method need only readily available measurements (for example, as obtainable using available vibration or control signals) in lieu of special "damage detectors." Given a proven damage tracking method, one acquires the ability to test available physics-based models or develop appropriate empirical models. Thus, the problem of failure prognosis can be reduced to a recursive filter design problem for estimating remaining time to failure.

In previous work [1,2] we have described a novel damage tracking method developed in a dynamical systems framework for studying damage evolution [3]. This method was shown to overcome most of the limitations discussed above. In particular, it was demonstrated both experimentally and analytically that the tracking output was in one-to-one relationship with the actual damage variable, and, indeed, under circumstances to be expected in many cases, not only one-to-one but also *linear*. Here we demonstrate an application and extension of this method to the development of a true damage diagnosis and prognosis algorithm.

3 Damage Tracking Based on Short-Time Dynamics

We view the state of a machine or component with damage as a point evolving in an extended phase space of a hierarchical dynamical system. This system consists of a "fast-time" subsystem coupled to a "slow-time" subsystem:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \boldsymbol{\mu}(\boldsymbol{\phi})), \tag{1a}$$

$$\dot{\boldsymbol{\phi}} = \epsilon \mathbf{g}(\boldsymbol{\phi}, \mathbf{x}), \tag{1b}$$

where: $\mathbf{x} \in \mathbb{R}^n$ is a directly observable, fast-time variable; $\boldsymbol{\phi} \in \mathbb{R}^m$ is a slow-time variable representing the damage that is "hidden" (i.e., not directly accessible); $\boldsymbol{\mu} \in \mathbb{R}^s$ is a function of $\boldsymbol{\phi}$ representing the material parameters in Eq. (1a); a rate constant $0 < \epsilon \ll 1$ defines a time scale separation between fast and slow-time dynamics; and overdots denote differentiation with respect to time t. Please note that for $\epsilon = 0$ the parameter vector $\boldsymbol{\mu}$ is a constant and system Eq. (1a) is stationary, and for $\epsilon \neq 0$ Eq. (1a) is nonstationary due to the evolution of $\boldsymbol{\phi}$. However, over the time scales of $\mathcal{O}(\epsilon)$ we consider Eq. (1a) to be quasistationary since drifts in $\boldsymbol{\mu}$ are negligible.

This formulation is appropriate for systems where hidden processes evolve on a much slower time scale than the directly ob-

servable dynamics that can be measured in real time, as is appropriate for systems with evolving damage. For example, crack growth in a spinning shaft can be characterized by a time scale of hours, days, weeks or even years, while a time scale of shaft vibration signature is characterized by milliseconds or seconds. In what follows we summarize the main ideas behind the damage tracking procedure [1], with a formulation that emphasizes its algorithmic implementation. We also note improvements on the procedure presented in the original papers.

3.1 The Damage Tracking Function and Reference Model
In an experimental context we usually do not have access to the analytical form of governing differential Eqs. (1). However, we have access to measurements from the fast-time system Eq. (1a). These measurements are usually, as will be assumed here, sampled at uniform time intervals t_s and the dynamics (or, more generally, the equivalent topological structure of the phase space) of Eq. (1a) can be reconstructed using a delay coordinate embedding [26,27]. In this procedure, the measured scalar time series $\{x(r)\}_{r=1}^M$ is converted to a series of vectors $\mathbf{y}^T(r) = [x(r), x(r+\tau), \ldots, x(r+(d-1)\tau)] \in \mathbb{R}^d$, where τ (multiple of t_s) is a suitable delay time and d is the appropriate embedding dimension. Embedding parameters τ and d are usually determined using the first minimum of the average mutual information [28] and method of false nearest neighbors [29], respectively.

The reconstructed state vectors are governed by an as yet unknown map of the form

$$\mathbf{y}(r+1) = \mathbf{P}(\mathbf{y}(r); \boldsymbol{\phi}). \tag{2}$$

In [1,2], the following *damage tracking function*

$$\mathcal{E}_k(r, \boldsymbol{\phi}) = \mathbf{P}^k(\mathbf{y}(r), \boldsymbol{\phi}) - \mathbf{P}^k(\mathbf{y}(r), \boldsymbol{\phi}_0) \tag{3}$$

was proposed. In Eq. (3), $\mathbf{P}^k$ is the k-th iterate of the map defined in Eq. (2). To actually calculate the tracking function $\mathcal{E}_k(r, \boldsymbol{\phi})$ for any given initial condition $\mathbf{y}(r)$, we need to know how the fast subsystem evolves over the time interval kt_s for the current, approximately constant, value of $\boldsymbol{\phi}$, as well as how this subsystem *would have* evolved for the reference value of $\boldsymbol{\phi}_0$. Since the system's fast time behavior for the current value of $\boldsymbol{\phi}$ is directly measurable (i.e., we can reconstruct the fast-time trajectory using a sensor measurement from the fast subsystem), the strategy is to compare it to the predictions of a *reference model* describing the fast system's behavior for $\boldsymbol{\phi} = \boldsymbol{\phi}_0$.

For fixed $\mathbf{y}(r)$, the tracking function can be expanded in a Taylor series in $\boldsymbol{\phi}$. For $\boldsymbol{\phi}$ sufficiently close to $\boldsymbol{\phi}_0$, it is shown in [1] that, assuming *linear observability* (i.e., assuming the first derivative of $\mathbf{P}^k$ with respect to $\boldsymbol{\phi}$ has maximal rank), the relationship between the tracking function and the damage variable can be well-approximated by an affine map

$$\mathcal{E}_k(r, \boldsymbol{\phi}) = \mathbf{C}(r)\boldsymbol{\phi} + \mathbf{c}(r), \tag{4}$$

where the matrix $\mathbf{C} = \partial \mathbf{P}^k / \partial \boldsymbol{\phi}$ is evaluated at $\boldsymbol{\phi} = \boldsymbol{\phi}_0$, and $\mathbf{c} = -\mathbf{C}\boldsymbol{\phi}_0$ is a column vector. Thus, under the above assumptions, the tracking function can be used to provide a linear measurement of, and therefore to track, the damage variable $\boldsymbol{\phi}$.[1] In [1,2], a suitably averaged (see section 3.2) $\mathcal{E}_k(r, \boldsymbol{\phi})$ was successfully used to track a scalar damage variable.

The reference model $\mathbf{P}(\cdot; \boldsymbol{\phi}_0)$ in Eq. (3) can be estimated in a variety of ways. In this work, local linear models are used as the simplest form of a globally nonlinear reference model

$$\mathbf{y}(r+1) = \mathbf{A}(r)\mathbf{y}(r) + \mathbf{a}(r), \tag{5}$$

where $\mathbf{A}(r)$ is an $n \times n$ matrix and $\mathbf{a}(r)$ is an $n \times 1$ vector. Equation (5) approximates Eq. (2) for a particular point $\mathbf{y}(r)$ in the reference system's reconstructed phase space. In practical appli-

---

[1]Note that tracking can be accomplished without necessarily knowing the matrices $\mathbf{C}$ and $\mathbf{c}$. In many applications, it is sufficient to know that the relationship is linear, without need for prior calibration to determine the "sensitivity" and "bias" of the measurements.
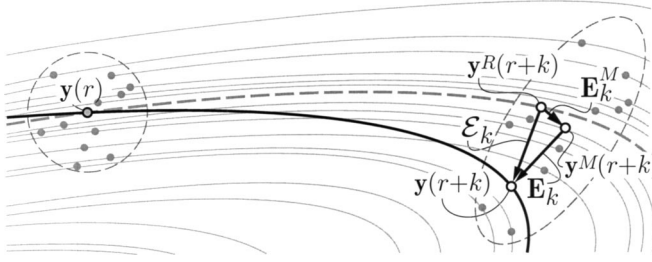
Fig. 1 Damage tracking function estimation. Solid black line is the current trajectory of the fast subsystem. Dashed gray line is the corresponding reference trajectory. Model is based on the reference trajectory points shown in gray.

cations, other modeling approaches, such as neural nets, radial basis functions, or auto regressive moving averages, may be more appropriate, either in terms of accuracy or implementability.

The parameters of the local linear models are determined by calculating the best linear fit between N nearest neighbors of $\mathbf{y}(r)$ and their future states for data taken in the reference condition. Then the damage tracking function (Eq. 3) for the initial point $\mathbf{y}(r)$ can be written as

$$\mathcal{E}_k(r,\boldsymbol{\phi}) = \mathbf{y}(r+k) - \mathbf{A}^k\mathbf{y}(r) - \mathbf{a}^k + \mathbf{E}^M(r) = \mathbf{E}_k(r,\boldsymbol{\phi}) + \mathbf{E}^M(r),$$
(6)

where, for simplicity, we have suppressed the dependency of $\mathbf{A}$ and $\mathbf{a}$ on r. In the above equation, $\mathbf{E}^M(r)$ represents the local linear model error and

$$\mathbf{E}_k(r,\boldsymbol{\phi}) = \mathbf{y}(r+k) - \mathbf{A}^k\mathbf{y}(r) - \mathbf{a}^k$$
(7)

is the estimated tracking function that can be determined experimentally (refer to Fig. 1). The use of $\mathbf{E}_k(r,\boldsymbol{\phi})$ in place of $\mathcal{E}_k(r,\boldsymbol{\phi})$ is justified if $\|\mathbf{E}^M\|$ is small compared to $\|\mathbf{E}_k\|$. In fact, as we discuss in the next section, the situation is somewhat better than this, since $\mathbf{E}_k(r,\boldsymbol{\phi})$ can be used to provide an *estimate* of $\mathcal{E}_k(r,\boldsymbol{\phi})$ by means of a suitably selected recursive filter.

3.2 Recursive Estimation of a Scalar Tracking Metric. In this paper we focus on scalar damage variables. This suggests that, instead of the vector tracking function defined in the previous section, we use $\|\mathcal{E}_k(r,\boldsymbol{\phi})\|$. However, as discussed in detail in [1], there will in general be fluctuations in the tracking function caused by two main sources that are not related to changes in the damage ϕ: (1) changes in the model fit error $\mathbf{E}^M$ from point to point in phase space; and (2) changes in the actual mapping of Eq. (4), also from point to point. We compensate for both sources of fluctuation by using a suitable filter, as described below. This approach has the added benefit of using all of the data available within one data record, and thus making the estimates more robust.

We use the data from an entire data record by considering the following *tracking metric*

$$e_k = \langle \|\mathcal{E}_k(r,\phi)\| \rangle,$$
(8)

where $\langle \cdot \rangle$ represents the *root mean square* (RMS) value over the index r (i.e., over an entire data record of time span t_D).

We can then attempt to estimate $\mathcal{E}_k(r,\phi)$ using only the measurable $\|\mathbf{E}_k(r,\phi)\|$ time series with an appropriate filter $\mathcal{F}$, so that Eq. (8) becomes:

$$e_k = \langle \mathcal{F}(\|\mathbf{E}_k(r,\phi)\|) \rangle.$$
(9)

In our previous work [1,2], the filter $\mathcal{F}$ was taken to be a weighted average in which the weights were proportional to the probability density of data points in the neighborhood of the current state. Here we improve the approach by defining a recursive filter, which is easier to implement and results in much better performance.

One way to design such a filter would be to make use of an explicit analytical model for the fast subsystem Eq. (1a), which could be used to derive a model describing the dynamics of $\|\mathbf{E}_k(r,\phi)\|$. In our case, we have assumed that such a model is not available. However, considering that the fast sampling time is very short ($t_s \ll t_D \ll 1/\epsilon$), a constant model should provide an acceptable approximation for the tracking function evolution over the sampling time interval t_s. That is, over the relatively short sampling time, the best estimate of the next "error state." $\|\mathbf{E}_k(r+1,\phi)\|$, is its current value, $\|\mathbf{E}_k(r,\phi)\|$, which is, obviously, a linear relationship.

For linear systems, the Kalman filter [30] provides the optimal minimum mean square error estimator, and it can be calculated recursively. Specifically, for $\mathcal{F}$ we consider a Kalman filter designed for random constant parameter estimation. Thus, we use the following linear process difference equations for the scalar tracking function $E(r) \equiv \|\mathcal{E}_k(r,\phi)\|$ and for the corresponding scalar measurement function $z(r) \equiv \|\mathbf{E}_k(r,\phi)\|$:

$$E(r+1) = E(r) + w(r+1)$$
$$z(r+1) = E(r+1) + v(r+1)$$
(10)

where the process (model) noise $w(r)$ and measurement noise $v(r)$ are assumed to be white, independent random variables with Gaussian distributions (i.e., $p(w) \sim N(0,Q)$ and $p(v(r)) \sim N(0,R(r))$).

In this case, the constant Q corresponds to the average amplitude of fluctuations in $\|\mathbf{E}_k(r)\|$ as the location in phase space changes along a given trajectory. In addition, $R(r)$ corresponds to fluctuations in $\|\mathbf{E}_k(r)\|$ due to changes in the local linear model accuracy from point to point in phase space. As in [1,2], we take $R(r) \propto d_r^D$, where d_r is the distance from the point $\mathbf{y}(r)$ to the farthest of all N nearest neighbors used for local modeling, and D is the average pointwise dimension [31]. The argument behind this choice is that the accuracy of the local linear models is proportional to the generalized volume occupied by the fixed number of nearest neighbors, and that the volume scales as d_r raised to the D power.

Note that the Kalman filter assumes that $w(l)$ and $v(l)$ have normal distributions. However, statistical characteristics of the tracking function output most likely will not be Gaussian since they are an outcome of a nonlinear process. In practice, very few processes possess noise terms that are normally distributed and, in case this deviation is due to some nonlinear measurement function, one usually tries to find an inverse transform to get Gaussian variables. If such a transform can not be found, however, Kalman filtering is still used since it provides accurate estimates of the first two moments of these distributions.

Finally, for each record, the damage tracking metric e_k and its variance are estimated as,

$$e_k = \langle E(r) \rangle, \quad \text{and} \quad \sigma_{e_k}^2 = \langle [E(r) - e_k]^2 \rangle.$$
(11)

4 Damage Diagnosis and Prognosis

Cusumano and Chatterjee [3] proposed that, given the hierarchical system of Eq. (1), the form of the damage evolution "law" of most use for predicting remaining useful life would be that related to the system of Eq. (1) by application of the concept of averaging. That is, the damage model used for failure prognosis is related to the original slow subsystem Eq. (1b) by taking the long time average of the vector field g along the solution to the $\epsilon = 0$ (i.e., the undamaged) system.[2]

This results, in principle, in a slow flow equation for the damage variable that is independent of the fast variable $\mathbf{x}$:

$$\dot{\phi} = \bar{g}(\phi),$$
(12)

---

[2]Since here we consider only scalar damage variables, g is also scalar-valued.

where the derivatives are taken with respect to "slow time" $\zeta = \epsilon t$ and $\bar{g}$ is the appropriately averaged version of the original g. In [2], this claim was examined using a mathematical model of a specific system having the general form of Eq. (1). In that particular case, the damage model provided by the averaged system Eq. (12) could be approximated analytically, and it was demonstrated that the output of the damage tracking algorithm followed its solutions very closely.

4.1 Recursive Estimation of Damage State. In general, reliable first-principles models for damage evolution are not available. In principle, the tracking methodology presented here permits the establishment of appropriate empirical damage evolution laws. However, in this paper we do not go through this procedure. We assume instead that the form of the damage model is known from previous calculations or is determined a posteriori from prior applications of the tracking algorithm. As mentioned in section 3.1, and demonstrated in earlier work, the tracking metric can be used to monitor an evolving damage variable. However, given a damage model, the tracking metric time series, which consists of one set of data of the form Eq. (11) for each data record, can be used to obtain an actual estimate of the damage state, together with an estimate of the uncertainty.

To construct a recursive estimator, we consider the tracking metric values to be measurements $z_\phi(r) \equiv e_k(r)$ that are linearly related to the damage state, following the linear observability assumption of Eq. (4). However, the damage process equation is usually nonlinear. This results in a set of estimator difference equations of the form:

$$\phi(r+1) = \phi(r) + \epsilon g(\phi(r))\Delta t + w_\phi(r+1),$$
$$z_\phi(r+1) = C\phi(r+1) + c + v_\phi(r+1),$$
(13)

where g is a given nonlinear function; $\Delta t = Mt_s = t_D$ for consecutive data records with no gaps; C and c are scalar parameters; and the process (model) noise $w_\phi(r)$ and measurement noise $v_\phi(r)$ are assumed to be white, independent random variables with Gaussian distributions $(p(w_\phi(r)) \sim N(0,Q_\phi(r))$ and $p(v_\phi(r)) \sim N(0,R_\phi(r)))$. Here, we take $R_\phi(r) \propto \sigma_{e_k}^2(r)$ to be the covariance associated with each measurement $z_\phi(r)$ and Q_ϕ is an estimate of the damage model error.

Given the nonlinearity of Eqs. (13), one must use a nonlinear technique to estimate the state of ϕ. While it is common practice to use the extended Kalman filter for nonlinear systems, it is not suitable for systems with strong nonlinearities. Here, we use so-called unscented filtering [32]. In the interest of completeness, we here summarize this approach as it applies to our problem.

Let us consider a scalar random variable $X(r)$ with mean $\bar{X}(r)$ and variance $P(r)$. We would like to predict the *a posteriori* mean $\bar{X}^-(r+1)$ and variance $P^-(r+1)$ of the random variable $X(r+1)$, where $X(r+1)$ is related to $X(r)$ by the nonlinear transformation $X(r+1) = G(X(r))$. In our case, the nonlinear transformation G includes all but the last term on the right hand side in the first of Eqs. (13). Then the general procedure at each iteration is:

1. Calculate the set of translated *sigma points* $\{\mathcal{X}_i(r)\}$ from the $P(r)$ as

$$\{\mathcal{X}_0(r), \mathcal{X}_1(r), \mathcal{X}_2(r)\} = \{\bar{X}(r), \bar{X}(r) - \sigma(r), \bar{X}(r) + \sigma(r)\},$$
(14)

where, $\sigma(r) = \sqrt{3P(r)}$.

2. The transformed set of sigma points are evaluated by $\mathcal{X}_i(r+1) = G(\mathcal{X}_i(r))$.

3. Compute $\bar{X}^-(r+1)$ and $P^-(r+1)$ by computing the mean and covariance of the set $\{\mathcal{X}_i(r+1)\}$.

4.2 Recursive Estimation of Time-to-Failure. Given a damage model Eq. (12) and a predefined failure surface (in the scalar case, simply a level), $\phi = \phi_F$, the time to failure t_F is
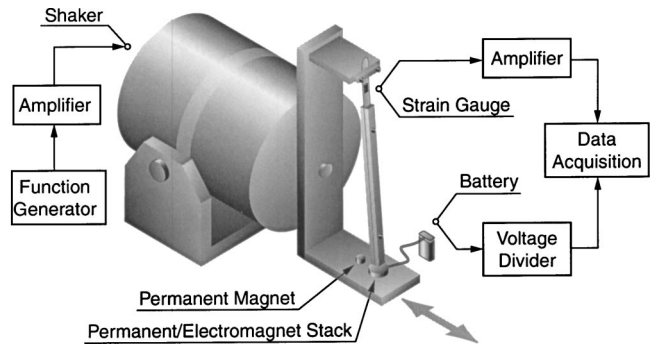

Fig. 2 Schematic diagram of the experimental apparatus of the two-well electro-mechanical oscillator.

uniquely defined by the solution to Eq. (12) with initial conditions $\phi(t_C)$ (the current damage state), such that $\phi(t_C + t_F) = 0$:

$$t_F = \int_\phi^{\phi_F} \frac{d\phi}{g(\phi)}.$$
(15)

Equation (15) can be evaluated analytically or numerically to give a finite t_F assuming that g is bounded away from zero.[3] (In the example studied in the next section, $g < 0$.)

Given the expression Eq. (15), damage state estimates can be used to generate derived measurements of the time to failure. A simple linear recursive estimator (Kalman filter) can then be used to obtain improved estimates of time to failure (together with uncertainty) using the following discrete-time state transition and measurement equations:

$$t(r+1) = t(r) - t_D + w(r),$$
$$t_F(r+1) = t(r+1) + v(r+1),$$
(16)

where $t(r)$ is the remaining time to failure, $t_F(r)$ is the current estimate (measurement) based on Eq. (15), t_D is the time length of a data record;[4] and $w(r)$ and $v(r)$ are assumed to be zero-mean white-noise processes with Gaussian distributions $(p(w) \sim N(0,Q_t(r))$ and $p(v(r)) \sim N(0,R_t(r)))$. Here, $R_\phi(r)$ is the covariance for each estimate of $\phi(r)$ (which is also obtained from the damage state estimator), and $R_t(r) = |(dt_F/d\phi)R_\phi(r)|$ is the covariance associated with each measurement.

5 Experimental Application to Electro-Mechanical System

The algorithm is tested experimentally using the system of [1,2], which is a modified version of the well known two-well magneto-mechanical oscillator [33]. The system consists of a single degree-of-freedom cantilever beam with an assembly of permanent magnets and an electromagnet near its end, providing a variable two-well potential (see Fig. 2). The electromagnet is powered by a standard 9 volt battery. A strain gauge is attached to the beam below the fixture and just above the stiffeners added to the beam to ensure a single mode vibrational response. The system is mounted on a shaker and is forced at about 6.3 Hz. The battery discharge process, which weakens the electromagnet strength, is viewed as a hidden damage process and the open circuit battery voltage is considered to be the damage variable.

The forcing amplitude was set to obtain nominally chaotic response at least in the beginning of experiment.[5] The experiment is

---

[3]This is equivalent to assuming that the damage model does not have an equilibrium point in the region of interest. See [3] for detailed discussion.

[4]$t_D = Mt_s$ in this case, but it can easily be generalized (to be irregular, for example).

[5]Note that the chaotic response is used in this case merely to simplify data acquisition: a more general method, such as stochastic interrogation [34] could be used for nonchaotic systems.
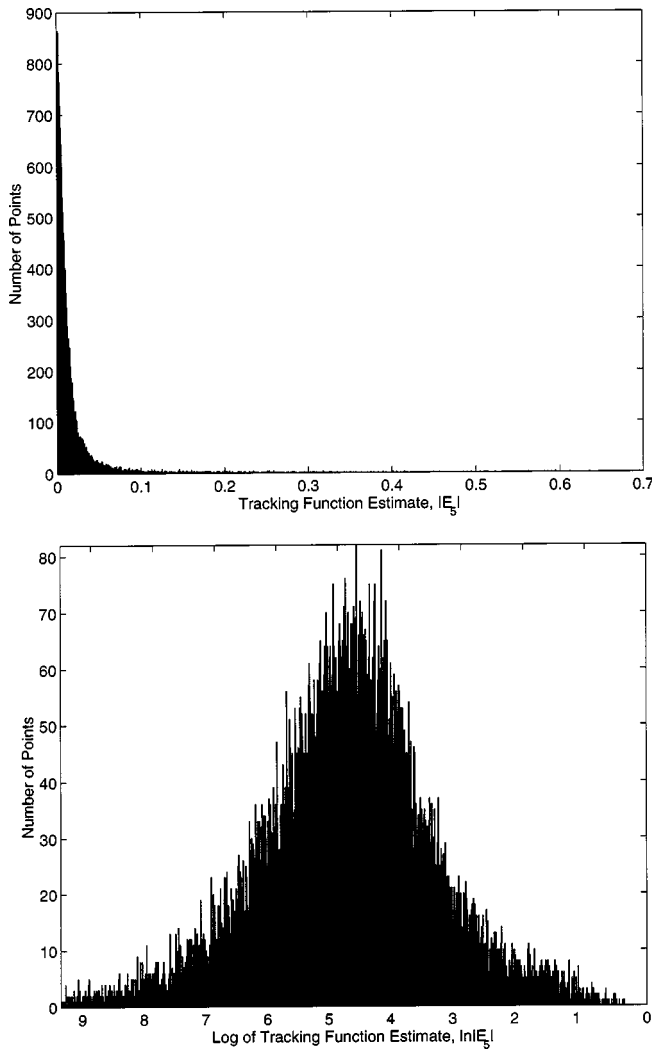
Fig. 3 Probability distributions of $|E_5|$ (upper) and $\ln|E_5|$ (lower). 2^{14} points were used for histograms.

started with a fresh 9 V battery and runs continuously for just under 7 hours. The stiffness in the electromagnet potential well decreases by a total of 6 percent due to the battery discharge. Strain gauge output is sampled at 100 Hz sampling frequency ($t_s = 0.01$ sec), digitized (12 bit A/D), and stored on a computer along with the independent voltage measurements taken from the electromagnets coil.

The system was chaotic throughout most of the experiment. However, numerous passages through windows of periodic behavior were also observed in the range of stiffnesses traversed by the system as the battery discharged. Delay time and embedding dimension for the reference data set were estimated to be $5t_s$ and 5, respectively. The first 2^{14} data points were used for the reference data set, and $N = 16$ nearest neighbors were used for the local linear model parameter estimation. The estimated average pointwise dimension of the reference data set was approximately $D = 2.82$.

After going through the embedding and modeling process, we estimated the tracking function e_5 by calculating the short-time prediction error $E_5(r)$, which is a scalar in this case, of the reference model along with the distance d_r between the point $\mathbf{y}(r)$ to the farthest nearest neighbor points $\mathbf{y}^{nn}(r)$ used in modeling. A sample distribution of $|E_5(r)|$ is highly asymmetrical and is shown on Fig. 3 (left). However, it can be approximated by a lognormal distribution. Indeed, if we take the natural log of the

absolute value of the tracking function we get an approximately Gaussian distribution (Fig. 3 right). Thus, to estimate the tracking function we apply the Kalman filter to $\ln(|E_5(r)|)$ and, then, transform it back using the exponential function. The filter parameters we used are: $R(l) = 10^3 d_r^D$ and $Q = 10^{-5}$.

A mathematical model for the experimental system was developed elsewhere by the authors [2]. It was shown that the following form of the battery voltage evolution law closely follows the experimental battery voltage data:

$$\dot{\phi} = -\epsilon(\phi - \psi)(1 + \gamma(\phi - \eta)^2), \qquad (17)$$

where $0 < \epsilon \ll 1$ is the rate constant, ψ represents residual voltage after battery discharge, η corresponds to the midpoint of the battery voltage operating range, and γ is a positive constant.

To determine the parameters of the battery voltage model Eq. (17), we make use of independently measured electromagnet coil voltage data. Since the recorded battery voltage signal has a modulation induced by the inductive coupling of the electromagnet electrical circuit and the beam oscillations, the local time average of the recorded voltage signal in each record was taken as an estimate of the open circuit battery voltage. Figure 4 (upper) uses thick gray line to depict the local mean of the measured battery voltage. On the same plot we present results of nonlinear curve fitting of Eq. (17) to the recorded voltage data using dashed black line. The nonlinear curve fitting resulted in the following parameters $\epsilon = 0.0429$, $\psi = 0.0766$, $\gamma = 4.2050$, and $\eta = 5.9168$.

The above model is used to formulate the discrete-time state transition equations for Eqs. (13). To determine the corresponding measurement equation we need to perform a calibration relating the battery voltage data and the estimated tracking function. Using a linear fit for Eq. (4) we have: $e_5 = C\phi + c$, where $C = -0.0085$, $c = 0.073$, and ϕ in this case corresponds to local time mean of the measured voltage data. This results in the following form of Eqs. (13):

$$\phi(r+1) = \phi(r) - \epsilon \Delta t(\phi(r) - \psi)(1 + \gamma(\phi(r) - \eta)^2) + w_\phi(r+1)$$
$$(18)$$

$$z_\phi(r+1) = C\phi(r+1) + c + v_\phi(r+1).$$

After obtaining the information needed to formulate the damage model above, experiments were conducted with consecutive data records of $M = 4{,}000$ points, with no overlapping sections or gaps between records. These records were used to calculate the mean $z_\phi(r) = e_5(r)$ and its variation $\sigma_{e_5}^2(r)$ using Eq. (11). Next, the nonlinear recursive filter for damage state estimation and Eq. (18) were used to estimate the battery voltage, as shown in Fig. 4 (lower) using a solid black line. In this case we used $R_\phi(r) = 10^2 \sigma_{e_5}^2(r)$ and $Q_\phi = 10^{-2}$.

By integrating Eq. (17) we can calculate the time elapsed from the initial to the current value of ϕ:

$$t(\phi) = -\frac{\ln\dfrac{\phi - \psi}{\sqrt{1 + \gamma(\phi - \eta)^2}} + (\eta - \psi)\sqrt{\gamma}\arctan[\sqrt{\gamma}(\phi - \eta)]}{\epsilon(1 + \gamma(\eta - \psi)^2)}$$
$$+ \frac{K}{\epsilon}, \qquad (19)$$

where K is a constant of integration.

We chose a failure value for the battery voltage of $\phi_F = 3.567$ V,[6] and used the following expression as the measurement input in Eqs. (16):

$$t_F(r) = t(\phi_F) - t(\phi(r)), \qquad (20)$$

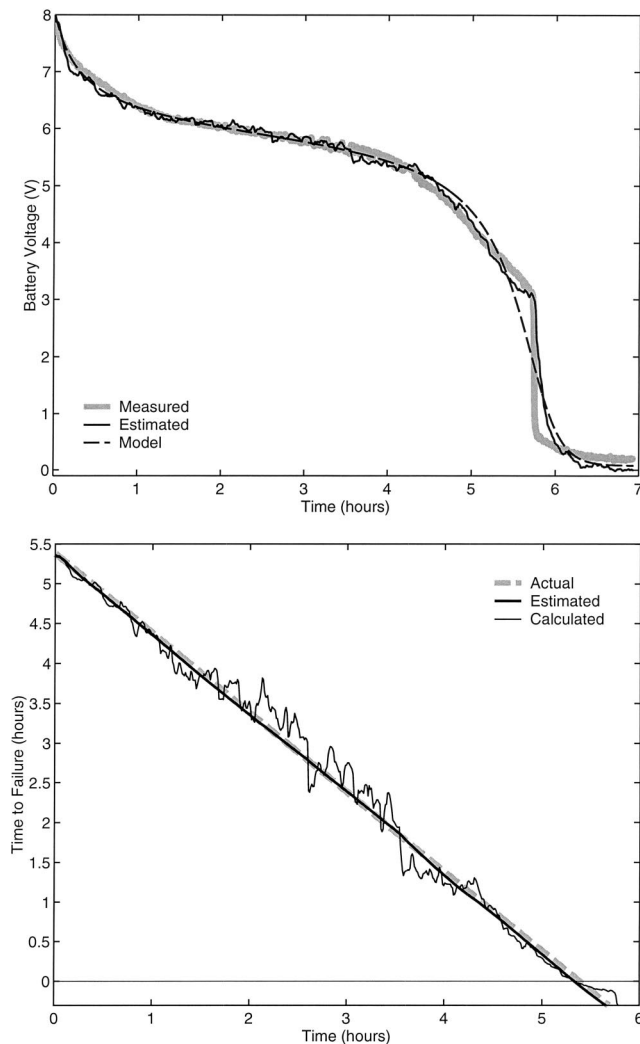[6]This value is arbitrary. Changing it would not significantly change our results.

Fig. 4 Damage state estimation and failure prognosis: (upper) plot of local mean of measured battery voltage (heavy gray line), fitted nonlinear battery discharge model (dashed black line), and recursively estimated battery state (solid black line) vs. time; (lower) time-to-failure predictions based on damage state estimates. In the time-to-failure predictions, the dashed heavy gray line indicates the true time to failure (known *a posteriori*), thin black line represents simple time-to-failure estimate using Eq. (20), and thick black line indicates the improved estimate using the failure prognostic filter of Eq. (16).

where $\phi(r)$ is a current estimate of the battery voltage. Results of the $t_F(r)$ calculation are shown in Fig. 4 (right) using a thin black line. To calculate corresponding variances we use the following equation:

$$R_F(r) \propto \left| \frac{dt_F(\phi(r))}{d\phi(r)} \right| \sigma_\phi^2(r) = \frac{\sigma_\phi^2(r)}{\epsilon |\phi(r) - \psi| (1 + \gamma(\phi(r) - \eta)^2)} \tag{21}$$

Results of this calculation are shown using a thick black line in Fig. 4 (right). For this calculation, we have used $Q_F = 10^{-4}$. The standard deviation for each estimate was too small to resolve in the plot. For a fixed ϕ_F, the failure prediction algorithm was able to provide accurate estimates of the time to failure (known *a posteriori* and shown by the heavy gray dashed line in Fig. 4) throughout the whole experiment.

6 Conclusions

In this paper we have presented a new procedure for damage diagnosis and prognosis. The algorithm is based on a multi-time-scale analysis of changes in the dynamics of a hierarchical system, and is not tied to any specific damage physics. It assumes that the system possesses time scale separation, i.e., damage process occurs on a much slower time scale than the observable dynamics of the system. It is assumed implicitly that the fast, directly observable subsystem is governed by some ordinary differential equation (Eq. 1*a*), however and explicit model need not be known. However, for failure prediction, a mathematical model for the damage evolution (Eq. 1*b*) is required.

We have described the major parts of our diagnostics/prognostics algorithm. In the first part, a tracking function based on reference model short-time prediction error statistics is used to obtain a tracking metric that provides a linear measurement of the current damage state. In the second part, we discussed how the tracking metric together with a damage evolution model are used to estimate current damage state more precisely. Finally, we described how predictions of the remaining time-to-failure can be obtained.

The algorithm was applied to an electro-mechanical oscillator with a drifting potential energy, caused by an electromagnet powered by a discharging battery. The battery discharge was considered to be a ''hidden'' damage process. As the battery discharged to complete failure, the system's stiffness in the potential energy well powered by the electromagnet declined by 6 percent. Throughout the experiment, the system underwent many bifurcations causing repeated periodic/chaotic transitions. Nevertheless, the method smoothly tracked the battery voltage, and predicted the remaining useful life of the battery well in advance of actual failure, using only strain measurements taken from the vibrating beam subsystem.

References

[1] Chelidze, D., Cusumano, J. P., and Chatterjee, A., 2002, ''Dynamical Systems Approach to Damage Evolution Tracking, Part 1: Description and Experimental Application,'' ASME J. Vibr. Acoust., **124**(2), pp. 250–257.
[2] Cusumano, J. P., Chelidze, D., and Chatterjee, A., 2002, ''Dynamical Systems Approach to Damage Evolution Tracking, Part 2: Model-Based Validation and Physical Interpretation,'' ASME J. Vibr. Acoust., **124**(2), pp. 258–264.
[3] Cusumano, J. P., and Chatterjee, A., 2000, ''Steps Towards a Qualitative Dynamics of Damage Evolution,'' Int. J. Solids Struct., **37**(44), pp. 6397–6417.
[4] Doebling, A. W., Farrar, C. R., Prime, M. B., and Shevitz, D. W., May 1996, ''Damage Identification and Health Monitoring of Structural and Mechanical Systems from Changes in Their Vibration Characteristics: A Literature Review,'' Technical Report LA-13070-MS, Los Alamos National Laboratory, Los Alamos, New Mexico 87545.
[5] Doebling, S. W., Farrar, C. R., and Prime, M. B., 1998, ''A Summary Review of Vibration-Based Damage Identification Methods,'' Sound Vib., **30**, pp. 91–105.
[6] Zou, Y., Tong, L., and P., S. G., 2000, ''Vibration-Based Model-Dependent Damage (Delimination) Identification and Health Monitoring for Composite Structures—A Review,'' J. Sound Vib., **230**(2), pp. 357–378.
[7] Qu, L., Xie, A., and Li, X., 1993, ''Study and Performance Evaluation of Some Nonlinear Diagnostic Methods for Large Rotating Machinery,'' J. Mech. Mach. Theory, **28**(5), pp. 699–713.
[8] Ma, J., and Li, C. J., 1995, ''On Gear Localized Defect Detection by Demodulation of Vibrations—A Comparison Study,'' E. Kannatey-Asibu, J., ed., *Proc. of Symp. on Mechatronics for Manufacturing in ASME International Mechanical Engineering Congress and Exposition*, MED-2-1, New York, NY, ASME, pp. 565–576.
[9] McFadden, P. D., and Smith, J. D., 1985, ''A Signal Processing Technique for Detecting Local Defects in a Gear from the Signal Averaging of Vibration,'' Proc. Inst. Mech. Eng., Part C: J. Mech. Eng. Sci., **199**(c4) ImechE-1985.
[10] Robert, T. S., and Lawrence, J. M., 1986, ''Detection, Diagnosis and Prognosis of Rotating Machinery,'' *Proc. of the 41st Meeting of the Mech. Failures Prevention Group*, Naval Air Test Center, Patuxent River, Maryland.
[11] McFadden, P. D., and Wang, W. J., 1993, ''Early Detection of Gear Failure by

Vibration Analysis—I. Calculation of the Time-Frequency Distribution,'' Mech. Syst. Signal Process., **37**(3), pp. 193–203.

[12] Liu, B., Ling, S., and Meng, Q., 1997, "Machinery Diagnisis Based on Wavelet Packets," J. Vib. Control, **3**, pp. 5–17.

[13] Baydar, N., Ball, A., and Kruger, U., 1999, "Detection of Incipient Tooth Defect in Helical Gears Using Principal Components," Starr, A. G., Leung, A. Y. T., Wright, J. R., and Sandoz, D. J., eds, *Integrating Dynamics, Condition Monitoring and Control for the 21st Century*, A. A. Balkema, Rotterdam, Brookfield, pp. 93–100.

[14] Isermann, R., 1984, "Process Fault Detection Based on Modeling and Estimation Methods—A Survey," Automatica, **20**(4), pp. 387–404.

[15] Gertler, J., 1988, "Survey of Model-Based Failure Detection and Isolation in Complex Plants," IEEE Control Syst. Mag., **8**(6), pp. 3–11.

[16] Natke, H. G., and Campel, C., 1997, *Model-Aided Diagnosis of Mechanical Systems: Fundamentals, Detection, Localization, Assesment*, Springer-Verlag, Berlin.

[17] Nijmeijer, H., and Fossen, T. I., 1999, "New Directions in Nonlinear Observer Design," *Lecture Notes in Control and Information Sciences*, Springer-Verlag, London, pp. 351–466.

[18] Wang, W., and Wong, A. K., 2002, "Autoregresive Model-Based Gear Fault Diagnosis," ASME J. Vibr. Acoust., **124**(2), pp. 172–179.

[19] McGhee, J., Henderson, I. A., and Baird, A., 1997, "Neural Networks Applied for the Identification and Fault Diagnosis of Process Valves and Actuators," J. Int. Measr. Confederation, **20**(4), pp. 267–275.

[20] Alessandri, A., and Parisini, T., 1997, "Model-Based Fault Diagnosis Using Nonlinear Estimators: A Neural Approach," *Proceedings of the American Control Conference*, Vol. 2, pp. 903–907.

[21] Li, C. J., and Fan, Y., 1997, "Recurrent Neural Networks for Fault Diagnosis and Severity Assessment of a Screw Compressor," *DSC ASME Dynamic Systems & Control Division*, Vol. 61, pp. 257–264.

[22] Jeon, Y. C., and Li, C. J., 1995, "Non-linear ARX Model Based Kullback Index for Fault Detection of a Screw Compressor," Mech. Syst. Signal Process., **9**(4), pp. 341–358.

[23] Li, Y., Billington, S., Zhang, C., Kurfess, T., Danyluk, S., and Liang, S. Y., 1999, "Adaptive Prognostics for Rolling Element Bearing Condition," Mech. Syst. Signal Process., **13**, pp. 103–113.

[24] Li, Y., Kurfess, T. R., and Liang, S. Y., 2000, "Stochastic Prognostics for Rolling Element Bearings," Mech. Syst. Signal Process., **14**(5), pp. 747–762.

[25] Swanson, D. C., Spencer, J. M., and Arzoumanian, S. H., 2000, "Prognostic Modelling of Crack Growth in a Tensioned Steel Band," Mech. Syst. Signal Process., **14**(5), pp. 789–803.

[26] Takens, F., 1981, "Detecting Strange Attractor in Turbulence," Rand, D. A. and Young, L. S., eds., *Dynamical Systems and Turbulence, Warwick, Springer Lecture Notes in Mathematics*, Springer-Verlag, Berlin, pp. 336–381.

[27] Sauer, T., Yorke, J. A., and Casdagli, M., 1991, "Embedology," J. Stat. Phys., **65**(3–4), pp. 579–616.

[28] Fraser, A. M., and Swinney, H. L., 1986, "Independent Coordinates for Strange Attractors from Mutual Information," Phys. Rev. A, **33**(2), pp. 1134–1140.

[29] Kennel, M. B., Brown, R., and Abarbanel, H. D. I., 1992, "Determining Embedding Dimension for Phase-Space Reconstruction Using a Geometric Construction," Phys. Rev. A, **45**(6), pp. 3403–3411.

[30] Grewal, M. S., and Angus, P. A., 1993, *Kalman Filtering/Theory and Practice*, Prentice Hall, Englewood Cliffs, New Jersey.

[31] Ott, E., 1993, *Chaos in Dynamical Systems*, Cambridge, New York.

[32] Julier, S. J., and Uhlmann, J. K., 1997, "New Extension of Kalman Filter to Nonlinear Systems," *Signal Processing, Sensor Fusion, and Target Recognition VI*, 3068, Proc. SPIE, pp. 182–193.

[33] Moon, F. C., and Holmes, P., 1979, "A Magnetoelastic Strange Attractor," J. Sound Vib., **65**(2), pp. 275–296.

[34] Cusumano, J. P., and Kimble, B., 1995, "A Stochastic Interrogation Method for Experimental Measurements of Global Dynamics and Basin Evolution: Application to a Two-Well Oscillator," Nonlinear Dyn., **8**, pp. 213–235.