 Asian Journal of Control, Vol. 3, No. 4, pp. 319-333, December 2001

EVOLUTIONARY-PROGRAMMING-BASED KALMAN FILTER FOR DISCRETE-TIME NONLINEAR UNCERTAIN SYSTEMS

Shu-Mei Guo, Leang-San Shieh, Ching-Fang Lin
and Norman P. Coleman

ABSTRACT

Some observations and improvements on the conventional Kalman filtering scheme to function properly are presented. The improvements can be achieved using the minimal principle evolutionary programming (EP) technique. A new linearization methodology is presented to obtain the exact linear models of a class of discrete-time nonlinear time-invariant systems at operating states of interest, so that the conventional Kalman filter can work for the nonlinear stochastic systems. Furthermore, a Kalman innovation filtering algorithm and such an algorithm based on the evolutionary programming optimal-search technique are proposed in this paper for discrete-time time-invariant nonlinear stochastic systems with unknown-but-bounded plant uncertainties and noise uncertainties to find a practically implementable "best" Kalman filter. The worst-case realization of the discrete-time nonlinear stochastic uncertain systems represented by the interval form with respect to the implemented "best" nominal filter is also found in this paper for demonstrating the effectiveness of the proposed filtering scheme.

KeyWords: Evolutionary programming, Kalman filter, nonlinear systems.

I. INTRODUCTION

Many uncertain issues related to system modeling, such as parameter variation and modeling error, generally result in uncertain mathematical models for most engineering systems or processes, to which the classical Kalman filter (*KF*) algorithm is generally not applicable. Therefore, one resorts to a variant of the conventional filtering scheme, the named robust Kalman filtering algorithm, which becomes more and more important in practice.

Recently, several approaches to robust *KF* have been discussed, based on, for instance, the criterion [22,29], uncertain system analysis [28,30], or set-valued estimation [4,12,21]. These modified versions of the *KF* computational scheme have an ability of handling uncertainties, but at the price of sacrificing the original means of optimality such as the linear unbiased property or the minimum statistical covariance assumptions or cost function, thereby actually reformulating and then solving a literally different estimation problem, so as to bypass the inherent difficulty of the embedded system uncertainties of the original KF algorithm.

An interval Kalman filtering (IKF) algorithm was proposed for an uncertain system described by interval matrices [3]. This IKF scheme, under exactly the same assumptions, can achieve exactly the same optimality (linear, unbiased, with minimum estimation error covariance). Moreover, it has exactly the same prediction-correction iterative structure. Most important of all, it is rigorous without additional conditions and approximations. The main problem with the IKF algorithm is its conservative property, due however to the conservative property of the interval mathematics and interval system modeling but not to the algorithm itself. Nevertheless, further improvement of the IKF scheme is expected and desirable.

Evolutionary programming (EP) was evolved and developed from the idea of genetic algorithms (GAs) [9, 14,15,19] and is a parallel optimization and computational

---

Manuscript received June 2, 2000; accepted November 30 2000.

Shu-Mei Guo is with Department of Computer Science and Information Engineering, National Cheng-Kung University, Tainan, Taiwan 701, R.O.C.

Leang-San Shieh is with Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77204-4005, U.S.A.

Ching-Fang Lin is with American GNC Corporation, 888 E. Easy Street, Simi Valley, CA 93065, U.S.A.

Norman P. Coleman is with U. S. Army Armament Center, Dover, NJ 07801, U.S.A.

This work was supported in part by the U. S. Army research office under Grant DAAG-55-98- 0198.

technique. As compared to the GAs, which use symbolic strings to describe a problem, EP uses functional forms [7, 8] so is more flexible and more suitable for solving complex engineering problems. It also follows the same competition principle like the GAs to eliminate unwanted candidates while preserving good ones; it also uses the same operations such as reproduction and mutation to gradually approach the global optima [20,23]. Taking advantage of EP for global optimization, the EP-based KF does not require more mathematical analysis than the IKF, and also preserves the same optimality and iterative computational structure as the classical KF under the same conditions.

In this paper some observations on the conventional KF, such as effects of the eigenspectrum of the system matrix and the excitation of output measurement, are given in Sec. 2. Then, the improvement of the Kalman filtering is newly proposed in Sec. 3, so that the KF scheme can work properly for linear time invariant stochastic systems. For further extensions of the above-mentioned improved KF to work properly from linear time invariant systems to some class of nominal nonlinear stochastic systems, an optimal linearization methodology is first proposed in Sec. 4 to obtain linear models of a class of nominal nonlinear systems at operating states of interest, using the newly proposed evolutionary programming (EP): the minimal principle approach. The proposed optimal linearization methodology yields the exact local linear models at operating states of interest and the optimal local linear models in the neighborhood of operating states of interest. Finally, based on another newly proposed minimal-maximal principle of EP, a novel EP-based Kalman filtering scheme is proposed in Sec. 5 to construct the "best" nominal Kalman filter for discrete-time time-invariant nonlinear stochastic uncertain systems. The worst-case realization of the discrete-time time-invariant nonlinear stochastic uncertain system with respect to the determined "best" filter is also given in Sec. 5. Kalman filtering has been widely used in many areas of industrial and government applications such as video and laser tracking systems, satellite navigation, ballistic missile trajectory estimation, radar, and fire control [5,17,25]. With the recent development of high-speed computers, the Kalman filter has become more useful even for very complicated real-time applications. The proposed design methodology enhances real-time applications of the Kalman filtering.

II. SOME OBSERVATIONS ON THE KALMAN FILTER

Consider a linear discrete multivariable system in state-space form

$$x(k+1) = Ax(k) + w(k), \tag{1}$$

$$y(k) = Cx(k) + v(k), \tag{2}$$

where $x(k)$ is an $n \times 1$ state vector, and $y(k)$ is an $p \times 1$ output vector with A and C being system matrices with appropriate dimensions. The vector $w(k)$ is the process noise due to disturbances and modeling inaccuracies and is assumed to be Gaussian, zero-mean $\bar{w}(k) = E[w(k)] = 0$ and white with the covariance $E\left|w(k)w^T(j)\right| = Q\delta(k-j)$, where $\delta(k-j) = I$ (identity matrix) when $k = j$; otherwise, $\delta(k-j) = 0$. The vector $v(k)$ is the measurement noise due to sensor inaccuracy with the same properties as $w(k)$ but has a different covariance matrix $E[v(k)v^T(k)] = R\delta(k-j)$. The sequences $w(k)$ and $v(k)$ are also assumed stationary and independent (orthogonal) of each other, i.e., $E[w(k)v^T(j)] = 0$ for any steps k and j. It is assumed that $x(0)$ is known in the form of its mean value $\bar{x}(0)$ and covariance $P(0)$.

Let the estimator have the form

$$\hat{x}(k+1 \mid k) = A\hat{x}(k \mid k-1) + K(k)[y(k) - C\hat{x}(k \mid k-1)]. \tag{3}$$

The reconstruction error $\tilde{x} = x - \hat{x}$ is governed by

$$\begin{aligned}\tilde{x}(k+1) &= A\tilde{x}(k) + w(k) - K(k)[y(k) - C\hat{x}(k \mid k-1)] \\ &= (A - K(k)C)\tilde{x}(k) + w(k) - K(k)v(k).\end{aligned} \tag{4}$$

The criterion is to minimize the variance of the estimation error, which is denoted by $P(k)$.

$$P(k) = E[(\tilde{x}(k) - E[\tilde{x}(k)])(\tilde{x}(k) - E[\tilde{x}(k)])^T]. \tag{5}$$

The mean value of $\tilde{x}$ is obtained from Eq. (4).

$$E[\tilde{x}(k+1)] = (A - K(k)C)E[\tilde{x}(k)] \tag{6}$$

The mean value of the reconstruction error is zero for all time $k \geq 0$ independent of $K(k)$ by assuming $E[\hat{x}(0)] = E[x(0)]$. However, it by no means implies that the estimation error $\tilde{x}(k)$ is white. Any random process governed by a known dynamics cannot be white, in general. Equations (4) and (5) now give

$$\begin{aligned}P(k+1) &= E[\tilde{x}(k+1)\tilde{x}^T(k+1)] \\ &= (A - K(k)C)P(k)(A - K(k)C)^T + Q + K(k)RK^T(k).\end{aligned} \tag{7}$$

After some mathematical manipulations [1], one has

$$K(k) = AP(k)C^T(R + CP(k)C^T)^{-1}, \tag{8}$$

$$\begin{aligned}P(k+1) &= AP(k)A^T \\ &\quad + Q - AP(k)C^T(R + CP(k)C^T)^{-1}CP(k)A^T.\end{aligned} \tag{9}$$

The reconstruction defined by Eqs. (3), (7), and (8) is called the Kalman filter. Notice that the criterion for choosing $K(k)$ is also to minimize the expected value of squared norm of $\tilde{x}(k)$, i.e., the length of the estimation error [13],

$$J_K(k) = E[\tilde{x}^T(k)\tilde{x}(k)] = trace\ \{E[\tilde{x}(k)\tilde{x}^T(k)]\} = trace\ [P(k)]. \tag{10}$$

The predictor (3) has the property that the state at time k is reconstructed from $y(k-1), y(k-2), \cdots$. It is also possible to derive the filter, which also uses $y(k)$, to estimate $x(k)$. The filter problem is solved by [1]

$$\hat{x}(k+1 \mid k+1) = A\hat{x}(k \mid k) + K(k+1)[y(k+1) - CA\hat{x}(k \mid k)], \tag{11}$$

where

$$K(k) = P(k \mid k-1)C^T[R + CP(k \mid k-1)C^T]^{-1}, \tag{12}$$

$$P(k \mid k-1) = AP(k-1 \mid k-1)A^T + Q, \tag{13}$$

$$P(k \mid k) = (I - K(k)C)P(k \mid k-1), \tag{14}$$

$$P(0 \mid 0) = P_0.$$

The notation $P(k \mid k-1)$ is used here instead of $P(k)$ to specify the available data; $P(k \mid k)$ can be interpreted as the covariance of the estimation error at time k given $Y_k = \{y(i) \mid i \le k\}$.

Some observations on the Kalman filter reconstructed by Eqs. (11)-(14), which are intended to find the best (or optimal) $\hat{x}(k)$ for the noisy stochastic (not deterministic) state $x(k)$ so that the variance of estimation error is as small as possible, are shown in the following.

If all eigenvalues of the system matrix A are small values, then the Kalman filtering scheme acts as a perfect "filter", which significantly filters out noise so that the estimated system state $\hat{x}(k)$ approaches the noise-free system state $\bar{x}(k)$ as close as possible. Essentially, a properly functioning Kalman filter is a low-pass filter with time-varying gain $K(k)$. It therefore possesses both noise rejection and smoothing properties [16]. The eigenvalue distribution of $A - K(k)C$ governs the "mean value" of the reconstruction error (6), and the eigenvalue distribution of $I - K(k)C$ governs the convergence index of the variance of the estimation error (14). The eigenspectrum $\sigma(I - K(k)C) = \{1, \ldots, 1\}$ indicates $P(k \mid k)$ converges to $P(k \mid k-1)$, which means the Kalman filter scheme converges; however, it by no means implies that the Kalman filter scheme converges to a desired condition. The criterion to minimize the variance of the estimation error (5) by no means implies that the best value of the state estimation error $\tilde{x}(k) = x(k) - \hat{x}(k)$ is zero, due to the stochastic property. However, it does mean that the smaller value of $P(k \mid k)$ is the better one, under the pre-required assumption that the Kalman filter functions properly. The properly functioning Kalman filter, shown in Example 1, yields an acceptable $E[\tilde{x}(k)]$ and $J_K(k) \approx trace[P(k \mid k)]$. However, whenever the Kalman filter is not functioning properly, the relation Eq. (10) does not hold anymore, due to the over-excited or under-excited output measurements (to be shown later).

To show how the magnitude of the eigenspectrum of the system matrix A affects the function of the Kalman filtering scheme, a system with a relatively large eigenspectrum is given in the following example.

Example 1. Let a linear discrete system be given as

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = \begin{bmatrix} 1.0 & 1.6 \\ 0 & 0.9 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + \begin{bmatrix} w_1(k) \\ w_2(k) \end{bmatrix}, \tag{15a}$$

$$y(k) = \begin{bmatrix} 0.5 & 0 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + v(k), \tag{15b}$$

where $Q = diag(0.01, 0.01)$, $R = 0.1$ and $x(0) = [x_1(0)\ x_2(0)]^T$ has mean $E[x(0)] = [0.2\ \ 0.2]^T$

This example and other similar higher-dimensional systems illustrate that if eigenvalues of the system matrix A have relatively large values, then the Kalman filter scheme works as a "state estimator", so that the estimated state $\hat{x}(k)$ approaches the noisy system state $x(k)$ as close as possible. Repeating the same process for this example time after time shows that the variation ranges of $x(k)$, $\hat{x}(k)$, and $J_K(k)$, respectively, are much larger than the cases when the system matrix A has relatively smaller eigenvalues.

To see how the output matrix C affects the Kalman filtering scheme, let's consider the same system shown in Example 1, except for various different output matrices C's, respectively, as follows:

$$y(k) = [1.0 \quad 0.0]x(k) + v(k), \tag{16a}$$

$$y(k) = [0.01 \quad 0.0]x(k) + v(k), \tag{16b}$$

$$y(k) = [0.0 \quad 0.0]x(k) + v(k). \tag{16c}$$

Since there exists a trade-off between $\sigma(A - K(k)C)$ and $\sigma(I - K(k)C)$, simulation results show that the Kalman filtering scheme does not work properly as either filter or state estimator, with the following steady-sate (started from time step $k = 5$) values:

$K(k) = [0.6315, 0.1529]^T$, $trace[P(k \mid k)] = 0.0824$,

$\sigma(A - K(k)C) = \{0.6343 \pm 0.4171\ i\}$,

and $\sigma(I - K(k)C) = \{1.0000, 0.3685\}$.

Notice that the estimated system state $\hat{x}_1(k)$ diverges from the system state $x_1(k)$. Whenever, $\sigma(A)$ is small and $K(k)$ is small, this is good for both Eqs. (6) and (14); nevertheless, whenever $\sigma(A)$ is large and $K(k)$ is small (large), it is just good for Eq. (14) (Eq. (6)), but not for Eq. (6) (Eq. (14)).

Based on the above observations, one may wonder if we can appropriately weigh the measurable output signals; i.e., weigh the output matrix C, so that the Kalman filter can work properly as a filter/estimator? Of course, it can be done, since weighting the measurable output signal does not affect the given noisy system state $x(k)$. However, the relationship between $K(k)$ and C given in Eqs. (12)-(14) is nonlinear, and there exists a trade-off between Eqs. (6) and (14), which induces us to propose the minimal principle Evolutionary Programming to obtain the optimal weighting matrix for C so that the length of the estimation error $J_k(k_f)$ is minimized, as in the next section.

Comparisons between $P(k|k)$ and $J_K(k)$ for cases Eqs. (15b), (16a)-(16c), and another case

$$y(k) = [2.0 \quad 0.0]x(k) + v(k) \tag{16d}$$

are given, respectively, for k = 200 as

$$J_K(200) = 1.0117 \times 10^5,\ trace[P(200|200)]$$

$$= 0.0354 \text{ for } C = [2.0 \quad 0.0],$$

$$J_K(200) = 21.4432,\ trace[P(200|200)]$$

$$= 0.0824 \text{ for } C = [1.0 \quad 0.0],$$

$$J_K(200) = 0.3998,\ trace[P(200|200)]$$

$$= 0.2166 \text{ for } C = [0.5 \quad 0.0],$$

$$J_K(200) = 18.4752,\ trace\ [P(200|200)]$$

$$= 41.4474 \text{ for } C = [0.01 \quad 0.0],$$

$$J_K(200) = 60.8169,\ trace\ [P(200|200)]$$

$$= 221.8975 \text{ for } C = [0.0 \quad 0.0],$$

The above results show that the $trace[P(k|k)]$ can not appropriately indicate a true time response whenever the Kalman filter works under the over-excited (Eqs. (16a) and (16d)) or under-excited (Eqs. (16b) and (16c)) output measurements due to the inappropriate output matrix C. As a result, the relationship $trace[P(k|k)] = J_K(k)$ does not hold anymore under the above-mentioned cases.

Before we go to the next section, an observation on the eigenspectum of the Kalman filtering scheme to the state-space self-tuning control for stochastic systems is briefly described as follows.

From a design point of view, a linear multivariable stochastic system with unknown system parameters and unknown noise statistics is first reformulated in a state-space innovation form, or an auto-regressive moving average model with exogerous input (ARMAX) form, suitable for parameter identification and state estimation [24]. Then the standard recursive extended least-squares estimation algorithm [18] is used to identify the unknown parameters. As a result, the Kalman gain matrix and so the system state can be estimated without solving a Riccati matrix equation. Therea-fter, an advantage control law can be employed as the desired self-tuner, which is finally implemented using the estimated system states in the observer coordinates for state-feedback control of the original multivariable stochastic system [24]. Consequently, the Kalman filter is implemented along with a controller in the form

$$u(k) = -F_0(k)\hat{x}_0(k) + H_0(k)r(k),$$

where $\{\hat{x}_0(k)\}$ is the realization of a Kalman filtering state sequence and $\{r(k)\}$ is the reference orbit, with time-varying coefficient matrices $F_0(k)$ and $H_0(k)$, respectively. Therefore, the closed-loop eigenvalues can be well assigned to be suitably small to have the dead-beat tracking property.

III. IMPROVED KALMAN FILTER: AN EVOLUTIONARY PROGRAMMING APPROACH

It is rather unexpected to realize that for the Kalman filter to work properly the process noise $w(k)$ should excite all the states and the measurement noise $v(k)$ should corrupt all of the measurements (i.e. $R > 0$). However, it is not easy to quantify it due to the trade-off nonlinear relationship between $K(k)$ and C, and so it is between Eq. (6) and Eq. (14). For the Kalman filtering scheme to function properly, i.e., to be well-excited by the weighted innovation error

$$e(k) = (\xi C)x(k) - (\xi C)\hat{x}(k|k) + v(k) \tag{17}$$

where

$$\xi = \left\| \left[\underline{\xi}_{ij} \ \overline{\xi}_{ij} \right] \right\| \in \Re^{P \times P},\ i = 1, 2, \ldots, P,$$

$$\text{and } j = 1, 2, \ldots, P, \tag{18}$$

which represents a linear combination of low-bound and upper-bound percentage changes of measurable output signals, an evolutionary programming technique is first proposed to minimize the Objective Function (OF) score

$$OF := E[\tilde{x}(k_f)^T \tilde{x}(k_f)] \approx \frac{1}{k_f} \sum_{i=1}^{n} \sum_{k=1}^{k_f} \tilde{x}_i^2(k), \tag{19}$$

where k_f is the final time step of interest, and the Kalman filter is constructed based on Eqs. (11)-(14), except that C is weighted as ξC, as shown in the following.

Suppose that the natural numbers are expressed in the scale of notation with radix R, so that

$$n = a_0 + a_1 R + a_2 R^2 + \cdots + a_m R^m,\ 0 \le a_i \le R. \qquad (20)$$

Write the digits of these numbers in reverse order, preceded by a decimal point. This gives the number

$$\phi_R(n) = a_0 R^{-1} + a_1 R^{-2} + \cdots + a_m R^{-m-1}. \qquad (21)$$

Holton [10] extended the two-dimensional result of Ban Der Corput [27] to κ-dimensions, when $R_1, R_2, \ldots, R_\kappa$ are mutually coprime [11].

Since $\phi_R(n) < 1$, to satisfy this range, scaling any varying parameter (e.g., a real number ε from its range $\left[\underline{\varepsilon}\quad \overline{\varepsilon}\right]$ to $\left[0\quad 1\right]$ is required. Let the interval real ($\Im\Re$) matrix $X \in \Im\Re^{n\times m}$ be a set of degenerate real matrices defined by

$$X = [L, U] = \{[x_{ij}] \,\big|\, l_{ij} \le x_{ij} \le u_{ij}; 1 \le i \le n, 1 \le j \le m\}, \qquad (22)$$

where L and U are constant real matrices. We introduce the variable ε_{ij}, $0 \le \varepsilon_{ij} \le 1$ such that

$$x_{ij} = l_{ij} + \varepsilon_{ij}(u_{ij} - l_{ij}) \qquad (23)$$

and use the notation

$$\varepsilon = [\varepsilon_{11}, \cdots, \varepsilon_{1m}, \varepsilon_{21}, \cdots, \varepsilon_{2m}, \varepsilon_{n1}, \cdots, \varepsilon_{nm}].$$

Then the interval matrix X can be denoted as $X(\varepsilon)$. Let $\varepsilon_{11} = \phi_2(n)$, $\varepsilon_{12} = \phi_3(n)$, $\varepsilon_{13} = \phi_3(n)$, and so on, to construct the desired initial population of size N (e.g., $N = 50$).

Define the minimal and maximal principles, respectively, as follows:

Minimal principle: Search some x^* in the solution set x, so that the objective function (denoted by "*OF*") value $OF(x)$ is minimal.
Maximal principle: Search some x^* in the solution set x, so that the objective function (denoted by "*OF*") value $OF(x)$ is maximal.

The developed EP algorithm for the minimal or maximal principle is described as follows:

1) Based on the quasi-random sequence (QRS) [11], form an initial population $\boldsymbol{P}_0 = [\boldsymbol{P}_1, \boldsymbol{P}_2, \ldots, \boldsymbol{P}_N]$ of size N by initializing each κ-dimensional solution vector $\boldsymbol{P}_i$ (used as individual) in $\boldsymbol{S}$. Here, population means a set of parameters we are looking for.
2) Assign each $\boldsymbol{P}_i$, $i = 1, \ldots, N$, an objective function score. Arrange $\boldsymbol{P}_i$, $i = 1, \ldots, N$, in descending order, starting from the best one generated from the objective function score.
3) Assign each sorted $\boldsymbol{P}_i$, $i = 1, \ldots, N$, a fitness function (denoted "*FF*") score to weigh those high-quality individuals in the pool of individuals based on the obtained objective function scores: For the maximal principle, use

$$FF(OF(\boldsymbol{P}_i)) = \left(\frac{\overline{\beta} - \underline{\beta}}{\overline{OF(\boldsymbol{P}_i)} - \underline{OF(\boldsymbol{P}_i)}}\right)(OF(P_i) - \underline{OF(\boldsymbol{P}_i)}) + \underline{\beta}; \qquad (24)$$

for the minimal principle, use

$$FF(OF(\boldsymbol{P}_i)) = \left[\left(\frac{\overline{\beta} - \underline{\beta}}{\overline{OF(\boldsymbol{P}_i)} - \underline{OF(\boldsymbol{P}_i)}}\right)(OF(P_i) - \underline{OF(\boldsymbol{P}_i)}) + \underline{\beta}\right]^{-1} \qquad (25)$$

This function linearly maps the real-valued space $[\underline{OF(\boldsymbol{P}_i)}, \overline{OF(\boldsymbol{P}_i)}]$ to any appropriate specified space, $[\underline{\beta}, \overline{\beta}]$ (e.g., [1, 10]), where $\underline{\beta} > 0$, for weighting the objective function scores. Hence, the better an individual is , the higher the objective function score that it will have.
4) Calculate the probability function (*PF*) score of each $\boldsymbol{P}_i$, $i = 1, \ldots, N$, using the fitness function score:

$$PF(FF(\boldsymbol{P}_i)) := PF(\boldsymbol{P}_i) = \frac{FF(\boldsymbol{P}_i)}{\sum_{i=1}^{N} FF(\boldsymbol{P}_i)}. \qquad (26)$$

5) Mutate each $\boldsymbol{P}_i$, $i = 1, \ldots, N$, based on statistics to double the population size from N to $2N$; assign $\boldsymbol{P}_{i+N}$ the following value:

$$\boldsymbol{P}_{i+N,j} := \boldsymbol{P}_{i,j}(1 + \mathrm{sgn}(N(0,1))\gamma(1 - FP(\boldsymbol{P}_i)), \qquad (27)$$

where $\boldsymbol{P}_{i,j}$ is the jth element in the ith individual, $\boldsymbol{N}(\mu, \sigma^2)$ is the Gaussian random variable with mean μ and variance σ^2, γ is a weighting factor for the percentage change of $\boldsymbol{P}_{i,j}$, and sgn(·) is the standard sign function. Whenever $\boldsymbol{P}_{i+N,j} \notin [\underline{P}_j, \overline{P}_j]$, some modification is required:

$$\boldsymbol{P}_{i+N,j} := \begin{cases} \underline{\boldsymbol{P}}_j & if\ \boldsymbol{P}_{i+N,j} < \underline{\boldsymbol{P}}_j \\ \overline{\boldsymbol{P}}_j & if\ \boldsymbol{P}_{i+N,j} > \overline{\boldsymbol{P}}_j \end{cases}. \qquad (28)$$

Properly adjusting the weighting factor γ can possibly avoid the undesired situation $\boldsymbol{P}_{i+N,j} \notin [\underline{P}_j, \overline{P}_j]$. It is notable that γ heavily dominates the convergence rate of the EP.

6) Calculate the objective function score of each $\boldsymbol{P}_{i+N}$, i = 1, …, N. Rank the objective function scores of $\boldsymbol{P}_i$, i = 1, …, 2N. Record $\boldsymbol{P}_i$, i = 1, …, 2N, in descending order, starting from the best individual in the pool of the population. The first N individuals are selected for the next generation, in which the top one of each generation, denoted $\boldsymbol{P}^*_{g,i}$, always survives and is selected for the next generation. Whenever $\boldsymbol{P}^*_{g,i}$ is no longer the best during the evolutionary process, update it by the newly generated best one.
7) Tune γ in the following way, to further avoid the search be trapped into a local extreme:

$$\gamma := \begin{cases} \gamma & if \left| OF(P^*_{g-1,i}) - OF(P^*_{g,i}) \right| > \eta \\ 1.5\gamma & if \left| OF(P^*_{g-2,i}) - OF(P^*_{g,i}) \right| \le \eta \\ 0.5\gamma & if \left| OF(P^*_{g-2,i}) - OF(P^*_{g,i}) \right| \le \eta \text{ and } \left| OF(P^*_{g-1,i}) - OF(P^*_{g,i}) \right| \le \eta \end{cases} \tag{29}$$

where η is some tolerable error bound and g is the generation index. Then, go to Step 2) and continue until the desired extreme value OF($\boldsymbol{P}^*_{g,i}$) cannot be further improved and/or the allowable generation is obtained. Then terminate the search process.

IV. OPTIMAL LINEARIZATION

Linearization such as Jacobian analysis is one of many useful techniques for analysis and design of nonlinear systems for local dynamic behavior [6]. The optimal linearization was first proposed by Teixeira and Zak [26] for continuous-time nonlinear systems followed by stabilizing controller design for uncertain nonlinear systems using fuzzy models. The proposed optimal linearization at the operating state, not necessarily the equilibrium state, yields the exact linear (not affine) model. Also it yields the optimal linear model defined by some convex constraint optimization criterion in the vicinity of the operating state. For linearization, Taylor expansion is also a common approach to use; however, a truncated Taylor expansion usually results in an affine rather than linear model due to the generally non-vanishing constant term. One exception is the trivial case where the equilibrium is zero, which, however, cannot be ensured throughout a nonlinear process. The objective of this section is to propose a new optional linearization method for nonlinear systems given in the discrete state-space form. Basically, the derivation in this section is similar to the one in Teixeira and Zak [26]; however, a further discussion on some relative topic of the linearized model, such as the observability, is also given in this section.

Consider the class of nonlinear systems described by

$$x(k+1) = f(x(k)), \tag{30}$$

$$y(k) = h(x(k)), \tag{31}$$

where $f : R^n \to \Re^n$ and $h : \Re^n \to \Re^n$ are nonlinear with continuous partial derivatives with respect to each of their variables at all steps k, where $x(k) \in \Re^n$ is the state vector at time index k, and $y(k) \in \Re^p$ is the measurable output vector at time index k. It is desired to have an exact local linear model ($A(k)$, $C(k)$) at an operating state of interest, $x(k) \in \Re^n$, in the form of

$$x(k+1) = A(k)x(k), \tag{32a}$$

$$y(k) = C(k)x(k), \tag{32b}$$

where $A(k)$ and $B(k)$ are constant matrices of appropriate dimensions. The linearization of the nonlinear system (30)-(31) is commonly represented by the truncated Taylor expansion as

$$x(k+1) - x_{eq}(k+1) = f(x_{eq}(k)) + A(k)[x(k) - x_{eq}(k)] \tag{33a}$$

or

$$x(k+1) = A(k)x(k) - A(k)x_{eq}(k), \tag{33b}$$

where $x_{eq}(k)$ is an equilibrium point. Clearly, this is an affine rather than linear model due to the generally non-vanished constant term in Eq. (33b). One exception is the trivial case where the equilibrium is zero, $x_{eq}(k) = 0$, which, however, cannot be ensured throughout a nonlinear control process. Suppose that we are given an operating state $x(k) \neq 0$; i.e., $x_i(k) \neq 0$ for i = 1, 2, …, n, which is not necessarily an equilibrium of the given system (30)-(31). The constraint $x(k) \neq 0$ will be released after the discussion on observability of the nonlinear system. The goal is also to construct an optimal local linear model, linear in x, such that in a neighborhood of $x(k)$, one has

$$f(x) \approx A(k)x, \tag{34}$$

$$h(x) \approx C(k)x \tag{35}$$

and

$$f(x) = A(k)x(k), \tag{36}$$

$$h(x) = C(k)x(k). \tag{37}$$

To satisfy these, let a_i^T denote the ith row of the matrix A (k), and represent Eq. (36) as

$$f_i(x) \approx a_i^T x, \; i = 1, 2, \ldots, n \tag{38}$$

and

$$f_i(x(k)) \approx a_i^T x(k),\ i = 1, 2, \ldots, n, \tag{39}$$

where $f_i : \Re^n \to \Re$ is the ith component of f. Then, expanding the left-hand side of Eq. (38) about $x(k)$ and neglecting the second and higher order terms, one has

$$f_i(x(k)) + [\nabla f_i(x(k))]^T (x - x(k)) \approx a_i^T x, \tag{40}$$

where $\nabla f_i(x(k)) : \Re^n \to \Re^n$ is the gradient column vector of f_i evaluated at $x(k)$. Due to Eq. (43), Equation (40) becomes

$$[\nabla f_i(x(k))]^T (x - x(k)) \approx a_i^T (x - x(k)), \tag{41}$$

in which x is arbitrary but should be "close" to $x(k)$ so that the approximation is good. To determine a constant vector, a_i^T, such that it is "as close as possible" to $[\nabla f_i(x(k))]^T$ and also satisfies $a_i^T x(k) = f_i(x(k))$, we may consider the following constrained minimization problem:

$$\min E := \frac{1}{2} \left\| \nabla f_i(x(k)) - a_i \right\|_2^2 \text{ subject to } a_i^T x(k) = f_i(x(k)). \tag{42}$$

Notice that this is a convex constrained optimization problem; therefore, the first order necessary condition for a minimum of E is also sufficient, which is

$$\nabla_{a_i} E + \lambda \nabla_{a_i} (a_i^T x(k) - f_i(x(k))) = 0, \tag{43}$$

$$a_i^T x(k) = f_i(x(k)), \tag{44}$$

where λ is the Lagrange multiplier and the subscript a_i in ∇_{a_i} indicates the gradient is taken with respect to a_i. It follows from Eq. (43) that

$$a_i - \nabla f_i(x(k)) + \lambda x(k) = 0. \tag{45}$$

Recall that we are studying the case where $x(k) \neq 0$, so by solving Eq. (45), we obtain

$$\lambda = \frac{x^T(k) \nabla f_i(x(k)) - f_i(x(k))}{\left\| x(k) \right\|_2^2}. \tag{46}$$

Substituting this into Eq. (45) gives

$$a_i = \nabla f_i(x(k)) + \frac{f_i(x(k)) - x^T(k) \nabla f_i(x(k))}{\left\| x(k) \right\|_2^2} x(k) \tag{47}$$

where $x(k) \neq 0$. Similar derivation can be applied to Eq. (35) to yield a similar result as

$$c_i = \nabla h_i(x(k)) + \frac{h_i(x(k)) - x^T(k) \nabla h_i(x(k))}{\left\| x(k) \right\|_2^2} x(k), \tag{48}$$

where c_i designates the ith row of the matrix $C(k)$. Note that, at an operating state of interest $x = x(k)$, the optimally linearized model ($f(x)$, $h(x)$) in Eqs. (36)-(37), which contains the optimal parameter matrices ($A(k)$, $C(k)$) obtained from the respective optimal parameter vector (a_i, c_i) in Eqs. (47)-(48), is identical to the exact nonlinear model ($f(x(k))$, $h(x(k))$) in Eqs. (30)-(31).

The observability matrix for the nonlinear system (30)-(31) is derived from the linearized model ($A(k)$, $C(k)$) (36)-(37), resulting in

$$O = \begin{bmatrix} \bar{C}(k) \\ \bar{C}(k)\bar{A}(k) \\ \bar{C}(k)\bar{A}^2(k) \\ \vdots \\ \bar{C}(k)\bar{A}^{n-1}(k) \end{bmatrix} \tag{49}$$

where $\bar{A}(k)$ and $\bar{C}(k)$ are constructed via the following rule: the ith columns of $A(k)$ and $C(k)$ are set to be zero whenever the ith component of $x(k)$ is zero.

The following example illustrates the aforementioned viewpoint.

Example 2. Based on the optimal linearization formula (47)-(48), the linear model of the following system [2]

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{bmatrix} = \begin{bmatrix} \alpha_1 e^{\alpha_2 x_1(k)} + \alpha_3 x_2(k) x_3(k) \\ \alpha_4 x_1(k) + \alpha_3 x_2(k) \\ 0.5 x_3(k) \end{bmatrix}, \tag{50a}$$

$$y(k) = \alpha_5 x_1^2(k) \tag{50b}$$

at any operating state $x(k)$ of interest is given by

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{bmatrix} = \begin{bmatrix} a_{11}(k) & a_{12}(k) & a_{13}(k) \\ \alpha_4 & \alpha_3 & 0 \\ 0 & 0 & 0.5 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix}, \tag{51a}$$

where

$$a_{11}(k) = \alpha_1 \alpha_2 e^{\alpha_2 x_1(k)} + [\alpha_1 (1 - \alpha_2 x_1(k))\ e^{\alpha_2 x_1(k)} - \alpha_3 x_2(k) x_3(k)] x_1(k) / \left\| x(k) \right\|_2^2,$$

$$a_{12}(k) = \alpha_3 x_3(k) + [\alpha_1 (1 - \alpha_2 x_1(k))\ e^{\alpha_2 x_1(k)} - \alpha_3 x_2(k) x_3(k)] x_2(k) / \left\| x(k) \right\|_2^2,$$

$$a_{13}(k)=\alpha_3 x_2(k)+[\,\alpha_1(1-\alpha_2 x_1(k))\;\; e^{\alpha_2 x_1(k)} \\ -\alpha_3 x_2(k)x_3(k)\,]x_3(k)\Big/\|x(k)\|_2^2,$$

and

$$y(k)=\Big[\,2\alpha_5 x_1(k)-\alpha_5 x_1^3(k)/\|x(k)\|_2^2-\alpha_5 x_1^2(k)x_2(k)/\|x(k)\|_2^2 \\ -\alpha_5 x_1^2(k)x_3(k)/\|x(k)\|_2^2\Big]x(k), \qquad (51b)$$

where α_i's, i = 1, 2, 3, 4 are some reasonable non-zero constant values. Let $x_i(k) \neq 0$ for i = 1, 2, 3. It is easy to show the rank of the observability matrix is full whenever $x_i(k) \neq 0$ for i = 1, 2, 3; i.e. $rank(O) = n = 3$, so it is observable. However, after few time steps,

$$x_3(k)=(0.5)^k x_3(0)\rightarrow 0.$$

When $x(k)$ is specified to be $x(k) = [x_1(k) \quad x_2(k) \quad 0]^T$, it results in

$$\begin{bmatrix} x_1(k+1)\\ x_2(k+1)\\ x_3(k+1)\end{bmatrix}=\begin{bmatrix}\overline{a_{11}}(k) & \overline{a_{12}}(k) & \alpha_3 x_2(k)\\ \alpha_4 & \alpha_3 & 0\\ 0 & 0 & 0.5\end{bmatrix}\begin{bmatrix} x_1(k)\\ x_2(k)\\ x_3(k)\end{bmatrix}:=A(k)x(k),$$

where

$$\overline{a_{11}}(k)=\alpha_1\alpha_2 e^{\alpha_2 x_1(k)}+[\,\alpha_1(1-\alpha_2 x_1(k))\;\; e^{\alpha_2 x_1(k)}\,]x_1(k)/\|x(k)\|_2^2$$

$$\overline{a_{12}}(k)=[\alpha_1(1-\alpha_2 x_1(k))e^{\alpha_2 x_1(k)}]x_2(k)/\|x(k)\|_2^2$$

$$y(k)=\Big[\,2\alpha_5 x_1(k)-\alpha_5 x_1^3(k)/\|x(k)\|_2^2,-\alpha_5 x_1^2(k)x_2(k)/\|x(k)\|_2^2,0\Big] \\ \times x(k):=C(k)x(k).$$

It is then straightforward to verify that the linearized system is observable since the $rank\lfloor(C^T(k), A^T(k)C^T(k), (A^T(k))^2C^T(k))^T\rfloor = n = 3$. However, directly substituting $x_3(k) = 0$ into the given system (50) yields

$$\begin{bmatrix} x_1(k+1)\\ x_2(k+1)\\ x_3(k+1)\end{bmatrix}=\begin{bmatrix}\alpha_1 e^{\alpha_2 x_1(k)}\\ \alpha_4 x_1(k)+\alpha_3 x_2(k)\\ 0\end{bmatrix}=\begin{bmatrix}\overline{a_{11}}(k) & \overline{a_{12}}(k) & 0\\ \alpha_4 & \alpha_3 & 0\\ 0 & 0 & 0\end{bmatrix}$$

$$\times\begin{bmatrix} x_1(k)\\ x_2(k)\\ x_3(k)\end{bmatrix}:=\bar{A}(k)x(k),$$

$$y(k)=\Big[\,2\alpha_5 x_1(k)-\alpha_5 x_1^3(k)/\|x(k)\|_2^2, \\ -\alpha_5 x_1^2(k)x_2(k)/\|x(k)\|_2^2,0\Big]x(k):=\bar{C}(k)x(k),$$

for convenience in checking the observability condition (wherever $x_3(k) = 0$). Here, $\bar{A}(k)$ and $\bar{C}(k)$ are constructed by replacing the third columns of $A(k)$ and $C(k)$ by zeros, respectively. Using $(\bar{A}(k), \bar{C}(k))$, it is easy to further verify that the given system does not belong to the class of observable systems.

Consequently, the constraint $x_i(k) \neq 0$ for $i = 1, 2, \ldots, n$ can be released provided that the matrices $(A(k), C(k))$ are replaced by $(\bar{A}(k), \bar{C}(k))$ for a Kalman filtering scheme and other design purposes, along with some decomposition technique in MatLab that can decompose the observable and unobservable portions.

V. EP-BASED KALMAN FILTERING SCHEME FOR UNCERTAIN NONLINEAR TIME-INVARIANT SYSTEMS

Consider the class of nominal discrete-time nonlinear time-invariant systems

$$\begin{cases} x(k+1)=f(x(k))+w(k)\\ y(k)=h(x(k))+v(k)\end{cases}$$

$$k = 0, 1, 2, \ldots, \qquad (52)$$

where $f(x(k))$ and $h(x(k))$ are $n \times 1$ and $p \times 1$ nonlinear vectors, and the noise sequences $\{w(k)\}$ and $\{v(k)\}$ satisfy the same assumptions as in model (1)-(2). Assume that both $f(x(k))$ and $h(x(k))$ are continuously differentiable with respect to each of their variables for all k. Then, when the system state $x(k)$ is available, we can apply the proposed optimal linearization method to obtain the following linear model

$$\begin{cases} x(k+1)=A(k)x(k)+w(k)\\ y(k)=C(k)x(k)+v(k)\end{cases}$$

$$k = 0, 1, 2, \ldots, \qquad (53)$$

where $x(k) \in \Re^n$ is the system state, $y(k) \in \Re^p$ is the measurement data, $A(k) \in \Re^{n\times n}$ and $C(k) \in \Re^{p\times n}$ are constant matrices obtained via the optimal linearization of nonlinear terms (36)-(37). Although $A(k)$ and $C(k)$ are time varying, it does not imply the given nonlinear system is time varying. In addition, assume that the system initial state $x(0)$ is a random vector independent of both $\{w(k)\}$ and $\{v(k)\}$, with $E\{x(0)\} = \bar{x}(0)$ and $\text{cov}\{x(0)\} = P(0) > 0$.

If all constant matrices $A(k)$ and $C(k)$ as well as all Q, R, $x(0)$, and $P(0)$ are known, then the classical Kalman filter algorithm is given by [2]

$$\begin{aligned} \hat{x}(k+1 \mid k+1) &= A(k)\hat{x}(k \mid k) + K(k+1)[y(k+1) \\ &\quad - C(k)A(k)\hat{x}(k \mid k)] \\ K(k) &= P(k \mid k-1)C^T(k)[R + C(k)P(k \mid k-1)C^T(k)]^{-1} \\ P(k \mid k-1) &= A(k)P(k-1 \mid k-1)A^T(k) + Q \\ P(k \mid k) &= [I - K(k)C(k)]P(k \mid k-1) \\ P(0 \mid 0) &= P_0 \end{aligned} \tag{54}$$

When system (52) has unknown-but-bounded uncertainties, it is described by an interval system of the form [3,4,30]

$$\begin{cases} x^I(k+1) = (A_0(k) + \Delta A(k))x^I(k) + w^I(k) \\ y^I(k) = (C_0(k) + \Delta C(k))x^I(k) + v^I(k) \end{cases}$$

$$k = 0, 1, 2, \ldots, \tag{55}$$

where $A_0(k)$ and $C_0(k)$ are matrices obtained via the proposed optimal linearization of the nominal nonlinear system, and $\Delta A(k)$ and $\Delta C(k)$ represent the bounded parameter uncertainties, while the noise $\{w^I(k)\}$ and $\{v^I(k)\}$ are the same mutually independent sequences as before, except that they now have interval covariance matrices

$$Q = Q_0 + \Delta Q \geq 0 \text{ amd } R = R_0 + \Delta R > 0.$$

In the following, for each realization of the interval system (55), i.e., a degenerate (i.e., real) nonlinear system actually appearing within the interval system (55), we use the following notation:

$$\begin{aligned} A_r(k) &\in A^I(k) := A_0(k) + \Delta A(k), C_r(k) \in C^I(k) \\ &:= C_0(k) + \Delta C(k), \\ Q_r &\in Q^I := Q_0 + \Delta Q, R_r \in R^I := R_0 + \Delta R, \\ K_r(k) &\in K^I(k), x_r(k) \in x^I(k), y_r(k) \in y^I(k), \hat{x}_r(k) \in \hat{x}^I(k). \end{aligned}$$

Thus, every realization satisfies the classical Kalman filter:

$$\begin{cases} \hat{x}_r(k+1 \mid k+1) = A_r(k)\hat{x}_r(k \mid k) + K_r(k+1)[y_r(k+1) - C_r(k)A_r(k)\hat{x}_r(k \mid k)] \\ K_r(k) = P_r(k \mid k-1)C_r^T(k)[R_r + C_r(k)P_r(k \mid k-1)C_r^T(k)]^{-1} \\ P_r(k \mid k-1) = A_r(k)P_r(k-1 \mid k-1)A_r^T(k) + Q_r \\ P_r(k \mid k) = [I - K_r(k)C_r(k)]P_r(k \mid k-1) \\ P_r(0 \mid 0) = P_0 \end{cases}, \tag{56}$$

where $y_r(k)$ is the measured output of the following realization

$$\begin{cases} x_r(k+1) = A_r(k)x_r(k) + w_r(k) \\ y_r(k) = C_r(k)x_r(k) + v_r(k) \end{cases}$$

$$k = 0, 1, 2, \ldots .$$

This framework will be used for the EP-based Kalman innovation filter to be further developed below, which requires the same conditions as the classical KF algorithm (54).

Let the interval Kalman filter be

$$\begin{aligned} \hat{x}^I(k+1 \mid k+1) &= A^I(k)\hat{x}^I(k \mid k) + K^I(k+1)[y^I(k+1) \\ &\quad - C^I(k)A^I(k)\hat{x}^I(k \mid k)], \end{aligned} \tag{57}$$

which contains every realization of Eq. (56) as the degenerate case. The main objective of this paper is to find the "best" nominal filter determined by some nominal nonlinear system represented by $(A_r^*(k), C_r^*(k), K_r^*(k))$, not necessarily the nominal filter determined by the nominal nonlinear system represented by $(A_0(k), C_0(k), K_0(k))$, such that the maximum filtering error $J_k(k_f)$ is minimized; namely, $J_{K_r^*}(k_f) = \min - \max J_k(k_f)$,
where

$$J_k(k_f) = J_k^{(1)}(k_f) := E\{(\tilde{x}_r(k_f) - E[\tilde{x}_r(k_f)])^T(\tilde{x}_r(k) - E[\tilde{x}_r(k_f)])\}$$

$$\text{if } E[\tilde{x}_r(k_f)] \not\to 0 \tag{58}$$

or

$$J_k(k_f) = J_k^{(2)}(k_f) := E[\tilde{x}_r^T(k_f)\tilde{x}_r(k_f)] \text{ if } E[\tilde{x}_r(k_f)] \to 0 \tag{59a}$$

$$\approx \frac{1}{k_f}\sum_{i=1}^{n}\sum_{k=1}^{k_f} \tilde{x}_i^2(k), \tag{59b}$$

where $\tilde{x}_r(k_f) = [\tilde{x}_1(k), \ldots, \tilde{x}_n(k_f)]^T$ and k_f is the final time step of interest. Here

$$\begin{cases} \tilde{x}_r(k) = x_r(k) - \hat{x}_{r,K_r^*}(k \mid k) \\ x_r(k+1) = f_r(x_r(k)) + w_r(k) \\ y_r(k) = h_r(x_r(k)) + v_r(k) \end{cases} \tag{60a}$$

and

$$\hat{x}_{r,K_r^*}(k+1 \mid k+1) = A_r^*(k)\hat{x}_{r,K_r^*}(k \mid k) + K_r^*(k+1)[y_r(k+1) - C_r^*(k)A_r^*(k)\hat{x}_{r,K_r^*}(k \mid k)], \quad (60b)$$

in which $\hat{x}_{r,K_r^*}(k \mid k)$ is the optimal estimate of the realization $x_r(k)$ when the practically implementable optimal Kalman filter ($A_r^*(k)$, $C_r^*(k)$, $K_r^*(k)$) is used based on the filtering algorithm (56). Notice that same $\{A_r^*(k), C_r^*(k)\}$ are used through Eq. (56) for the design phase of KF (see Level 1: Design Level); however, the respective $\{A_r(k), C_r(k)\}$ in Eq. (60a) are not equal to $\{A_r^*(k), C_r^*(k)\}$ in Eq. (56) for the test phase of the designed KF (see Level 2: Test Level), respec-tively, in general. Theoretically, when the Kalman filter (60) is replaced by Eq. (3), the explicit representation of $J_k^{(2)}(k+1)$ is given by

$$\begin{aligned} E[\tilde{x}(k+1)\tilde{x}^T(k+1)] &= E\{[(A_r^*(k) - K_r^*(k)C_r^*(k))\tilde{x}(k) \\ &\quad + (\Delta A_r(k) - K_r^*(k)\Delta C_r(k))x(k)] \\ &\quad \times [(A_r^*(k) - K_r^*(k)C_r^*(k))\tilde{x}(k) \\ &\quad + (\Delta A_r(k) - K_r^*(k)\Delta C_r(k))]^T\} \\ &\quad + K_r^* E[v(k)v^T(k)]K_r^{*T} + E[w(k)w^T(k)] \\ &= E\{[(A_r^*(k) - K_r^*(k)C_r^*(k))\tilde{x}(k) \\ &\quad + (\Delta A_r(k) - K_r^*(k)\Delta C_r(k))x(k)] \\ &\quad \times [A_r^*(k) - K_r^*(k)C_r^*(k)]\tilde{x}(k) \\ &\quad + (\Delta A_r(k) - K_r^*(k)\Delta C_r(k))x^T(k)\} \\ &\quad + K_r^*(R_r + \Delta R_r)K_r^{*T} + Q_r + \Delta Q_r, \quad (61) \end{aligned}$$

where $\Delta A_r(k) = A_r(k) - A_r^*(k)$ and $\Delta C_r(k) = C_r(k) - C_r^*(k)$. Nevertheless, it is a really difficult task to solve Eq. (61). A more complex explicit form can also be derived for the case of filter (60). Due to the complication for solving the above-mentioned explicit formulas, it is desired to replace Eq. (59a) by Eq. (59b). Some interpretation on the objective of this paper is further given as follows.

When system (52) has uncertainties described by an interval model, it takes on the form

$$\begin{cases} x^I(k+1) = f^I(x^I(k)) + w^I(k) \\ y^I(k) = h^I(x^I(k)) + v^I(k) \end{cases}$$

$$k = 0, 1, 2, \ldots, \quad (62)$$

where $f^I(x^I(k))$ and $h^I(x^I(k))$ are interval nonlinear vectors as defined before, with continuous partial derivatives with respect to each of their variables at all steps k. Suppose that the uncertainties and the dominant parts of the system nonlinearities can be confined into (and only into) the interval matrices $A^I(k)$ and $C^I(k)$. Under the assumption that $\hat{x}_r(k \mid k) = x_r(k)$, it yields $\hat{A}_r(k) = A_r(k)$ and $\hat{C}_r(k) = C_r(k)$, where $\{\hat{A}_r(k), \hat{C}_r(k)\}$ and $\{A_r(k), C_r(k)\}$ are matrices obtained via the $\hat{x}_r(k \mid k)$-based and $x_r(k)$-based optional linearizations of Eq. (62), respectively. Nevertheless, due to the fact that $\hat{x}_r(k \mid k) \neq x_r(k)$, in general, there always exist some perturbations between $\{\hat{A}_r(k), \hat{C}_r(k)\}$ and $\{A_r(k), C_r(k)\}$, denoted by $\Delta\hat{A}_r(k) = A_r(k) - \hat{A}_r(k)$ and $\Delta\hat{C}_r(k) = C_r(k) - \hat{C}_r(k)$, which also yield a similar representation to Eq. (55). Therefore, the aforementioned objective in this paper can be further extended to work for systems represented by Eqs. (55) to (62).

The proposed EP-based optimization process for finding the "best" nominal filter, among virtually infinitely many others in an interval system, is summarized and described as follows.

Consider the uncertain discrete-time nonlinear time-invariant system (62). The objective here is to find the practically implementable "best" nominal Kalman filter (60) yielding the desired min-max $J_{K_r^*}(k_f)$. The procedure of the desired EP-based design-test scheme is as follows.

Level 1. Design Level-Design the Filter

1) Generate a κ-dimensional initial population $\boldsymbol{P}$ of size N, denoted by $\boldsymbol{IP} = \{\boldsymbol{P}_{d,0,i};\ i = 1, 2, \ldots, N\}$, and a spare population of size N' (need not be equal to N), denoted by $\boldsymbol{SP} = \{\boldsymbol{P}_{d,0,i}:\ i = N+1, N+2, \ldots, N+N'\}$. Here, the index 0 is the initial generation index $g = 0$ and d indicates that the quantity is at the design level. This task is done by using QRS to initialize each individual $\boldsymbol{P}_{d,0,i} \in \boldsymbol{IP} \cup \boldsymbol{SP}$, for $i = 1, \ldots, N, N+1, \ldots, N+N'$.
2) Use the proposed optimal linearization formulas (47)-(48); except for the replacement of the unmeasurable $x_r(k)$ by the measurable $\hat{x}_r(k \mid k)$, to form the linear model of Eq. (62), and apply the classical KF scheme (56) to obtain the KF gain $K_r(k)$, so that the realized Kalman filter of each individual $\boldsymbol{P}_{d,0,i}$ is constructed based on Eq. (56).
3) Assign to each $\boldsymbol{P}_{d,0,i}$ an objective function (OF) score:

$$\max J_{KF(\cdot)}(k_f) = OF(\mathrm{P}_{t,g,i}^*; \mathrm{P}_{d,g,i}, K_{d,g,i}^*),$$

where the index t indicates that the quantity is at the test level (see Level 2 below). This OF can be the one defined in Eq. (58) or (59). By going through the test level (Level 2 described below), we can find the above maximal objective function value.
4) Receive the message from the test level about the degenerate Kalman filter KF ($\boldsymbol{P}_{d,g,i}$, $K_{d,g,i}$) so obtained, if it satisfies the stability requirement. If not, this matrix has to be replaced by one from the spare population $\boldsymbol{SP}$,

until stability is achieved.

5) Apply the minimal principle operator of EP to create a new population of higher quality. Go to Step 2) at this level and change the step index from generation $g = 0$ to $g = 1$. Continue the programming until the minimum value of max $\boldsymbol{J}_{KF(\cdot)}(k_f)$ is reached. This resulting stage will provide the associated "best'' nominal Kalman innovation filter $KF(\boldsymbol{P}^*_{d,g,i}, K^*_{d,g,i})$. At this stage, the corresponding min – max $\boldsymbol{J}_{KF(\cdot)}(k_f)$ cannot be further improved or the allowable tolerance is met).

Level 2. Test Level - Test the Designed Filter

1) Generate a κ-dimensional initial population of size M (need not be equal to N or N'), by using QRS to initialize each individual $\boldsymbol{P}_{t,g',i'} \in \boldsymbol{S}$, for $i' = 1, 2, \ldots, M$.
2) Assign to each $\boldsymbol{P}_{t,g',i'}$, $i' = 1, 2, \ldots, M$, an objective function (OF) score. This OF can be the one defined in Eq. (58) or Eq. (59). If some OF score is much higher than others, it means the Kalman filter being tested is infeasible. In this case, the stability is consequently not guaranteed, so send a message to Level 1 about this situation and then terminate the process at this level; otherwise, continue the process.
3) Apply the maximal principle operator of the EP to create a new population of higher quality.
4) Go to Step 2) at this level and repeat the steps, until the maximal value of $\boldsymbol{J}_{KF(\cdot)}(k_f)$ is reached. This resulting stage will provide the max $\boldsymbol{J}_{KF(\cdot)}(k_f)$ under the realization of the interval system in terms of $\boldsymbol{P}^*_{t,g',i'}$, which cannot be further improved (or the allowable tolerance is met).
5) Inform Level 1 about the finding of an individual with the highest quality at this level, $\boldsymbol{P}^*_{t,g',i'}$ (which is actually the worst-case of estimation error and will be minimized at Level 1, as discussed above).

VI. ILLUSTRATIVE EXAMPLE

Example 3. Consider the nominal discrete-time nonlinear time-invariant system [2]

$$\begin{cases} x(k+1) = f(x(k)) + w(k) \\ y(k) = h(x(k)) + v(k) \end{cases} \tag{63}$$

where

$$f(x(k)) = \begin{bmatrix} \alpha_{10} e^{\alpha_{20} x_1(k)} - 0.5 x_2(k) \\ -0.5 x_1(k) - \alpha_{30} x_3(k) \end{bmatrix} = \begin{bmatrix} 0.02 e^{-2.2 x_1(k)} - 0.5 x_2(k) \\ -0.5 x_1(k) - 0.5 x_3(k) \end{bmatrix}$$

$$h(x(k)) = \alpha_{40} x_1^2(k) := 1.0 x_1^2(k),$$

$Q_0 = diag(q_{11}, q_{22}) = diag(0.01, 0.01)$, and $R_0 = r_{11} = 0.1$. The optimal linear model of $f(x(k))$ and $h(x(k))$ is given, respectively, by

$$f(\hat{x}(k|k)) = A(k)\hat{x}(k|k) := \begin{bmatrix} a_{11}(k) & a_{12}(k) \\ a_{21}(k) & a_{22}(k) \end{bmatrix} \begin{bmatrix} \hat{x}_1(k|k) \\ \hat{x}_2(k|k) \end{bmatrix}$$

where

$$a_{11}(k) = \alpha_{10}\alpha_{20} e^{\alpha_{20}\hat{x}_1(k|k)} + \alpha_{10}[1 - \alpha_{20}\hat{x}_1(k|k)] e^{\alpha_{20}\hat{x}_1(k|k)} \hat{x}_1(k|k) / \|\hat{x}(k|k)\|_2^2$$

$$a_{12}(k) = \alpha_{10}[1 - \alpha_{20}\hat{x}_1(k|k)] e^{\alpha_{20}\hat{x}_1(k|k)} \hat{x}_2(k|k) / \|\hat{x}(k|k)\|_2^2$$

$$a_{21}(k) = -0.5\hat{x}_1(k|k),\ a_{22}(k) = \alpha_{30},$$

and

$$h(\hat{x}(k|k)) = C(k)\hat{x}(k|k)$$

where

$$C(k) = \Big[2\alpha_{40}\hat{x}_1(k|k) - \alpha_{40}\hat{x}_1^3(k|k) / \|\hat{x}(k|k)\|_2^2, \quad -\alpha_{40}\hat{x}_1^2(k|k)\hat{x}_2(k|k) / \|\hat{x}(k|k)\|_2^2 \Big].$$

Here, $\hat{x}(k|k)$ is the estimated system state based on the following Kalman filter scheme

$$\begin{aligned} \hat{x}(k+1|k+1) &= A(k)\hat{x}(k|k) + K(k+1)[y(k+1) - C(k)A(k)\hat{x}(k|k)] \\ K(k) &= P(k|k-1)C^T(k)[R + C(k)P(k|k-1)C^T(k)]^{-1} \\ P(k|k-1) &= A(k)P(k-1|k-1)A^T(k) + Q \\ P(k|k) &= [I - K(k)C(k)]P(k|k-1) \\ P(0|0) &= P_0. \end{aligned} \tag{64}$$

Let $x(0) = [0.2, 0.2]^T$, $\hat{x}(0|0) = [0.45 \ \ 0.25]^T$, and $P(0|0) = diag(0.01, 0.01)$. Simulation results based on the Kalman filtering scheme (63)-(64) are given by parts shown in Fig. 1, where $E[\tilde{x}(k_f)] = E[x(k_f) - \hat{x}(k_f|k_f)] = [-0.0096, 0.0031]^T$, and $J_k(k_f) = 0.0508$. Figure 1 shows that the above KF scheme provides the characteristics of over-excited innovation error, due to the inapposite output measurement $h(x(k))$. To overcome the above drawback, the EP-based improved KF scheme, newly proposed in

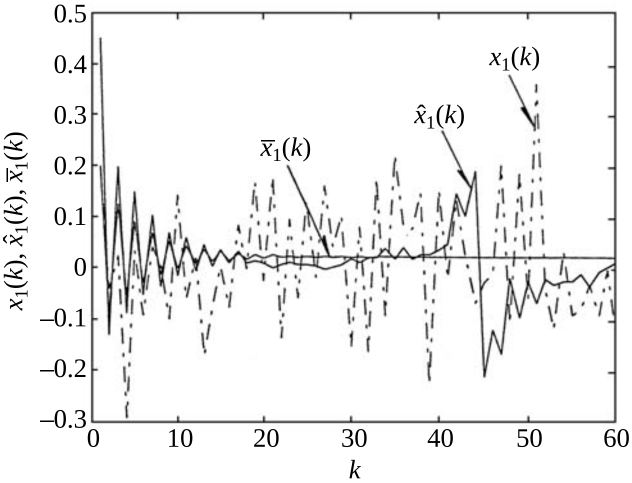

Fig. 1. State responses for $\alpha_4 = 1.0$ – $x_1(k)$: system state, $\hat{x}_1(k)$: estimated system state, and $\bar{x}_1(k)$: noise-free systems state.

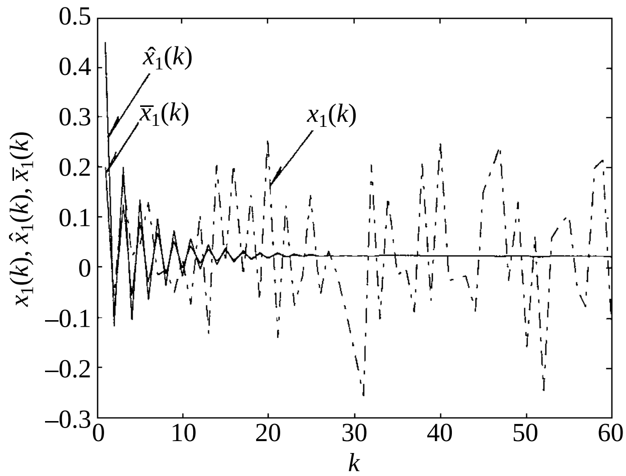

Fig. 2. State responses for $\alpha_4 = 0.0158$ – $x_1(k)$: system state, $\hat{x}_1(k)$: estimated system state, and $\bar{x}_1(k)$: noise-free systems state.

Sec. 3, is employed to yield the optimal output measurement $h^*(x(k)) = 0.0158x_1^2(k)$ with satisfied state responses where $E[\tilde{x}(k_f)] = [0.0004, -0.0071]^T$ and $J_k(k_f) = 0.0312$, given by parts shown in Fig. 2, in which the output measurement $h(x(k))$ is weighted by $h^I(x^I(k)) := (\xi^I\alpha_{40})x_1^2(k) = \alpha_4^I x_1^2(k)$, where the interval range of α_4^I is given by $\alpha_4^I = [1/200 \;\; 1]\; \alpha_{40} = [0.0050 \;\; 1.0000]$.

Next, consider the discrete-time nonlinear time-invariant system with unknown-but-bounded plant uncertainties and noise uncertainties

$$\begin{cases} x^I(k+1) = f^I(x^I(k)) + w^I(k) \\ y^I(k) = h^I(x^I(k)) + v^I(k) \end{cases} \tag{65}$$

where

$$f^I(x^I(k)) = \begin{bmatrix} \left[\,\underline{\alpha_1}\;\overline{\alpha_1}\,\right] e^{\left[\,\underline{\alpha_2}\;\overline{\alpha_2}\,\right] x_1(k)} - 0.5x_2(k) \\ -0.5x_1(k) + \left[\,\underline{\alpha_3}\;\overline{\alpha_3}\,\right] x_2(k) \end{bmatrix},$$

$$\underline{\alpha_1} = 0.9\alpha_{10},\; \overline{\alpha_1} = 1.1\alpha_{10},\; \underline{\alpha_2} = 0.9\alpha_{20},\; \overline{\alpha_2} = 1.1\alpha_{20},$$

$$\underline{\alpha_3} = 0.9\alpha_{30},\; \overline{\alpha_3} = 1.1\alpha_{30};$$

$$h^I(x^I(k)) = \left[\,\underline{\alpha_4}\;\overline{\alpha_4}\,\right] x_1^2(k),\; \underline{\alpha_4} = 0.9(0.0158),\; \overline{\alpha_4} = 1.1(0.0158);$$

$$Q^I = diag\,([\,\underline{\alpha_5}\;\overline{\alpha_5}\,]\,[\,\underline{\alpha_6}\;\overline{\alpha_6}\,]),\; \underline{\alpha_5} = \underline{\alpha_6} = 0.8q_{11},$$

$$\overline{\alpha_5} = \overline{\alpha_6} = 1.2q_{11};$$

$$R^I = \left[\,\underline{\alpha_7}\;\overline{\alpha_7}\,\right],\; \underline{\alpha_7} = 0.8r_{11},\; \overline{\alpha_7} = 1.2r_{11}.$$

It is desired to find a practically implementable "best" KF (60b) for any possible parameter set ($f_r(x_r(k))$, $h_r(x_r(k))$, $w_r(k)$, $v_r(k)$) of the uncertain system (65), so that the worst-case mean-square of the state estimation Eq. (58) or Eq. (59) is minimized. Also, the worst-case possible set (fr ($f_r(x_r(k))$, $h_r(x_r(k))$, $w_r(k)$, $v_r(k)$) of the uncertain system with respect to the implemented "best" nominal filter is designed to be determined for demonstrating the effectiveness of the proposed filtering scheme.

Following the design-test procedure described in the previous section, in which $\kappa = 7$, $N = N' = M = 50$, $\left|\underline{\beta},\, \bar{\beta}\right|$ = [1, 10], γ's for the minimal principle and the maximal principle are –1.2 and 0.8, respectively, the "best" Kalman filter is constructed based on the following scheme:

$$\hat{x}_{r,K_r^*}(k+1 \mid k+1) = A_r^*(k)\hat{x}_{r,K_r^*}(k \mid k) + K_r^*(k+1)[y_r(k+1) - C_r^*(k)A_r^*(k)\hat{x}_{r,K_r^*}(k \mid k)], \tag{66}$$

where $A_r^*(k)$, $C_r^*(k)$, and $K_r^*(k)$ are pre-computed (i.e. in an off-line way) based on the following realized nominal nonlinear system for the filter, denoted by $x_f(k+1)$ and filter pair:

$$system\text{: } x_f(k+1) = \begin{bmatrix} 0.0202e^{-2.2136x_{f,1}(k)} - 0.5x_{f,2}(k) \\ -0.5x_{f,1}(k) - 0.5468x_{f,2}(k) \end{bmatrix} + w_f(k) \tag{67}$$

$$y_f(k) = 0.0142x_{f,1}^2(k) + v_f(k),$$

where $Q_f = diag(0.0100, 0.0111)$, $R_f = 0.1156$, and $x_f(0) = [0.2, 0.2]^T$;

$$filter\text{: } \hat{x}_{f,K_r^*}(k+1 \mid k+1) = A_r^*(k)\hat{x}_{f,K_r^*}(k \mid k) + K_r^*(k+1)[y_f(k+1) - C_r^*(k)A_r^*(k)\hat{x}_{f,K_r^*}(k \mid k)] \tag{68}$$

where

$$A_r^*(k) = \begin{bmatrix} \alpha_1\alpha_2 e^{\alpha_2 \hat{x}_{f,K_r^*,1}(k|k)} + \alpha_1[1-\alpha_2\hat{x}_{f,K_r^*,1}(k|k)e^{\alpha_2\hat{x}_{f,K_r^*,1}(k|k)}\hat{x}_{f,K_r^*,1}(k|k)/\|x_{f,K_r^*}(k|k)\|_2^2, & \alpha_1[1-\alpha_2\hat{x}_{f,K_r^*,1}(k|k)e^{\alpha_2\hat{x}_{f,K_r^*,1}(k|k)}\hat{x}_{f,K_r^*,2}(k|k)/\|x_{f,K_r^*}(k|k)\|_2^2 \\ -0.5\hat{x}_{f,K_r^*,1}(k|k), & \alpha_3 \end{bmatrix}$$

$\alpha_1 = 0.0202$, $\alpha_2 = -2.2136$, $\alpha_3 = -0.5468$;

$$C_r^*(k) = \begin{bmatrix} 2\alpha_4\hat{x}_{f,K_r^*,1}(k|k) - \alpha_4\hat{x}^3_{f,K_r^*,1}(k|k)/\|\hat{x}_{f,K_r^*}(k|k)\|_2^2 \\ -\alpha_4\hat{x}^2_{f,K_r^*,1}(k|k)\hat{x}_{f,K_r^*,2}(k|k)/\|\hat{x}_{f,K_r^*}(k|k)\|_2^2 \end{bmatrix}^T,$$

$\alpha_4 = 0.0142$;

$$K_r^*(k) = P_r^*(k|k-1)C_r^{*T}(k)[R_r^* + C_r^*(k)P_r^*(k|k-1)C_r^{*T}(k)]^{-1}, \quad (69)$$

$$P_r^*(k|k-1) = A_r^*(k)P_r^*(k-1|k-1)A_r^{*T}(k) + Q_r^*,$$

$$P_r^*(k|k) = [I - K_r^*(k)C_r^*(k)]P_r^*(k|k-1),$$

$$P_r^*(0|0) = P_0^* = diag\,(0.01,\ 0.01),\ \hat{x}(0|0) = [0.45,\ 0.25]^T.$$

The $y_f(k+1)$ in Eq. (68) is any measured output of the realized interval nonlinear system (67). The worst-case realization $y_r(k)$ of Eq. (65) is also given by

$$x(k+1) = \begin{bmatrix} 0.0198e^{-2.3385x_1(k)} - 0.5x_2(k) \\ -0.5x_1(k) - 0.5380x_2(k) \end{bmatrix} + w(k), \quad (70)$$

$$y(k) = 0.0143x_1^2(k) + v(k),$$

where

$Q = diag(0.0091, 0.0083)$ and $R = 0.1129$.

Based on the proposed design-test procedure, some quite satisfactory simulation results on the realized worst-case system (70) to the obtained "best" Kalman filter (66)-(69) are shown by parts in Fig. 3.

VII. CONCLUSIONS

For the Kalman filtering scheme to work properly for both discrete-time linear and nonlinear systems, it is required to have well-excited innovation error. An evolutionary programming technique is first proposed in this paper to minimize the mean of squared state estimation errors at the final time step of interest. A new optimal linearization methodology is also presented in this paper for the above-mentioned design goal. Based on the analytically linearization model at each sampling time, the observability of the specific class of discrete-time nonlinear nominal systems is discussed in this paper. Furthermore, the evolutionary-programming-based Kalman filtering scheme for uncertain discrete-time nonlinear systems is newly proposed in this paper. The design-test procedure based *KF* scheme utilizes the global-search capability of EP to find the practically implementable "best" nominal filter for the discrete-time nonlinear uncertain system. The worst-case realization of the discrete-time nonlinear uncertain system represented by the interval form with respect to the "best" nominal filter is also given in this paper.

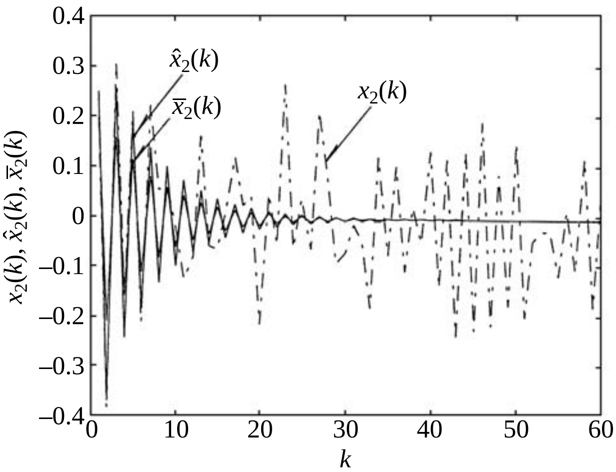

Fig. 3. State responses for $-x_2(k)$: system state, $x_2(k)$: estimated system state, and $x_2(k)$: noise-free systems state.

REFERENCES

1. Äström, K.J. and B. Wittenmark., *Computer Controlled Systems: Theory and Design*, Prentice-Hall, Englewood Cliffs, New Jersey, pp. 268-272 (1984).
2. Chen, G., G. Chen and S.H. Hsu, *Linear Stochastic Control Systems*, CRC Press, New York, pp. 174-197 (1995).
3. Chen, G., J.W. Wang and L.S. Shieh, "Interval Kalman Filtering,'' *IEEE Trans. Aerosp. Electron. Syst.*, Vol. 33, No. 1, pp. 250-259 (1997).
4. Chen, G., Q. Xie and L.S. Shieh, "Fuzzy Kalman Filtering," *J. Inform. Sci.*, Vol. 109, No. 3, pp. 197-209 (1998).
5. Chui, C.K. and G. Chen, *Kalman Filtering with Real-Time Applications*, Springer-Verlag, New York, pp. 50-66 (1987).
6. Dabney, J. B.and T.L. Harman, *Mastering SIMULINK*

2, Prentice-Hall, New Jersey, pp. 195-205 (1998).
7. Fogel, D.B., *Evolutionary Computation: The Fossil Record,* IEEE Press, Piscataway, New Jersey, pp. 3-14 (1998).
8. Fogel, D.B., "An Overview of Evolutionary Programming," *Evolutionary Algorithms,* L. D. Davis, K. D. Jong, M. D. Vose and L. D. Whitley Eds., Springer-Verlag, New York, IMA-Vol. 11, pp. 89-109 (1999).
9. Goldberg, D.E., *Genetic Algorithms in Search, Optimization and Machine Learning,* Addison-Wesley, Reading, MA, pp. 59-88 (1989).
10. Halton, J.H., "On the Efficiency of Certain Quasi-Random Sequences of Points in Evaluating Multidimensional Integrals," *Numerische Mathematik*, Vol. 2, pp. 84-90, with Corrigenda on p.196 (1960).
11. Hammersley, J. M and D.C. Handscomb, *Monte Carlo Methods,* Methuen, London, pp. 33-36 (1964).
12. Hong, L., "Disturbed Filtering Using Set Models for Systems with Non-Gaussian," *Approximate Kalman Filtering,* Ed. G. Chen, World Scientific Publishing Co., Singapore, pp. 139-160 (1993).
13. Juang, J. N., *Applied System Identificatio*n*,* Prentice-Hall, Englewood Cliffs, New Jersey, pp. 192-227 (1994).
14. Kramer, S.C. and R.C.IV. Martin, "Direct Optimization of Gain Scheduled Controllers Via Genetic Algorithms," *J. Guid. Contr. Dynam.*, Vol. 19, No. 3, pp. 636-642 (1996).
15. Krishnakumar, K and D.E. Goldberg, "Control System Optimization Using Genetic Algorithm," *J. Guid. Contr., Dynam.*, Vol. 15, No. 3, pp. 735-740 (1992).
16. Lewis, F.L., *Applied Optimal Control and Estimation: Digital Design and Implementation*, Prentice-Hall, Englewood Cliffs, New Jersey, pp. 491-493 (1992).
17. Lin, C.F., *Modern Navigation, Guidance, and Control Processing,* Prentice Hall, New Jersey, pp. 414-460 (1991).
18. Ljung, L., "Convergence Analysis of Parameter Identification Methods,'' *IEEE Trans. Automat. Contr.*, Vol. 23, No. 5, pp. 770-783 (1978).
19. Man, K.F., K.S. Tang and S. Kwang, *Genetic Algorithms: Concepts and Design*, Springer-Verlag, London, pp. 75-153 (1999).
20. Michalewicz, Z., *Genetic Algorithm + Data Structure = Evolution Programs*, Springer-Verlag, New York, pp. 159-177 (1996).
21. Morrel, D. and W.C. Stirling, "Distributed Kalman Filtering," *Approximate Kalman Filtering,* Ed. G. Chen, World Scientific Publishing Co., Singapore, pp. 139-160 (1993).
22. Nagpal K.M and P. Khargonekar, "Filtering and Smoothing in an Setting," *IEEE Trans. Automat. Contr.*, Vol. 36, No. 2, pp. 152-166 (1991).
23. Polak, E. and Y.Y. Wardi, "Nondifferentiable Optimization Algorithm for Designing Control Systems Having Singular Value Inequalities," *Automatic*a, Vol. 18, No. 3, pp. 267-283 (1982).
24. Shieh, L.S., Y.L. Bao and F.R. Chang, "State-Space Self-Tuning Regulators for General Multivariable Stochastic Systems,'' *IEE Proc., Pt. D*, Vol. 136, No. 1, pp. 17-27 (1989).
25. Sorenson, H.W., "Kalman Filtering Techniques," in *Advances in Control Systems Theory and Applications of Kalman Filtering*, Ed. C. T. Leondes, Vol. 3, Academic Press, New York, pp. 219-292 (1979).
26. Teixeira, M.C.M.and S.H. Zak, "Stabilizing Controller Design for Uncertain Nonlinear Systems Using Fuzzy Models," *IEEE Trans. Fuzzy Syst.*, Vol. 7, No. 2, pp. 133-142 (1999).
27. Van Der Corput, J. C., "Verteilungsfunktionen,'' *Proc. Kon Akad. Wet.,* Amsterdam, Vol. 38, pp. A13-A21, 1058-1066 (1935).
28. Xie, L., Y.C. Soh and C.M. Desouza, "Robust Kalman Filtering for Uncertain Discrete-Time System," *IEEE Trans. on Automat. Contr.*, Vol. 39, No. 6, pp. 1310-1314 (1994).
29. Xie, L.D., C.E. Souza and M Fu, "Estimation of Discrete-Time Linear Uncertain Systems," *Int. J. Robust Nonlin. Contr.,* Vol. 1, No. 1, pp. 111-123 (1991).
30. Zhang, X.F., A.W. Heemink and J.C.H. Van Eijkeren, "Performance Robustness Analysis of Kalman Filter for Liner Discrete-Time Systems under Plant and Noise Uncertainty," *Int. J. Syst. Sci.*, Vol. 26, No. 2, pp. 257-275 (1995).

Shu-Mei Guo received the B.S. degree in Department of Business Administration from National Chung-Hsing University, Taiwan, R.O.C. in 1982 and the M.S. degrees in Department of Computer and Information Science from New Jersey Institute of technology, U.S.A. in 1987. She received the Ph.D. degree in Computer and Systems Engineering from University of Houston, U.S.A. in May 2000. Since June 2000, she has been an assistant professor in Department of Computer System and Information Engineering, National Cheng-Kung University, Taiwan. Her research interests include various applications on evolutionary programming, chaos systems, Kalman filtering, fuzzy methodology, sampled-data systems, and computer and systems engineering.

Leang-San Shieh received his B.S. degree from the National Taiwan University, Taiwan in 1958, and his M.S. and Ph.D. degrees from the University of Houston, Houston, Texas, in 1968 and 1970, respectively, all in electrical engineering. He joined the Department of Electrical and Computer Engineering at the University of Houston in 1971 as an Assistant Professor, and he was promoted to Associate Professor in 1974 and Professor in 1978. Since 1988, he has been the Director of the Computer and Systems Engineering. Dr. Shieh is a member of IEEE and AIAA, and is a Registered Professional Engineer in the State of Texas. He was the recipient of more than ten College Outstanding Teacher Awards, the 1973 and 1997 College Teaching Excellence Awards, and the 1988 College Senior Faculty Research Excellence Award from the Cullen College of Engineering, University of Houston, and the 1976 University Teaching Excellence Award from the University of Houston. Further more, he received the Honor of Merit from Instituto Universitario Politecnico, Republic of Venezuela in 1978. His fields of interest are digital control, optimal control, self-tuning control and hybrid control of uncertain systems He has published more than two hundred articles in various scientific journals. He co-authored a research monograph, "An Algebraic Approach to Structural Analysis and Design of Multivariable Control Systems", (Springer, New York) in 1988.

Ching-Fang Lin is the founder and President of the American GNC (AGNC) Corporation and the driving force behind AGNC's many accomplishments. Dr. Lin not only heads the management team at AGNC, but he also dictates the technical direction of the company. He is responsible for numerous technical advances within the field of Guidance, Navigation, and Control (GNC). His career spans more than 25 years of teaching, research, industrial, and management experience in the realm of GNC.

He began his illustrious career at the University of Michigan in Ann Arbor, where he received his Ph.D. degree in Computer, Information, and Control Engineering in 1980. During his teaching career at the University of Michigan and at the University of Wisconsin in Madison, and as the founder and president of the American GNC Corporation, he has consistently pursued several areas of critical importance to modern technology and has carefully evaluated, in both theory and practice, the implications of their integration. He has authored over 350 technical publications and, from 1990 to 2000, Dr. Lin was responsible for over 100 patent invention disclosures. Over the same time period, he was responsible for over 500 government contract reports. In the last ten years, he led the effort to introduce over 30 GNC products. He is author of the book series includes Modern Navigation, Guidance, and Control Processing, Book II, and Advanced Control System Design, Book III. These two books are the best selling in their field, across the globe.

Norman P. Coleman received his B. A. degree in mathematics from the University of Virginia in 1965 and M.A and Ph.D. degrees from Vanderbilt University in 1967 and 1969, respectively. He is currently Chief of the Automation and Robotics Laboratory of the US Army Armament Research, Development and Engineering Center (ARDEC), Picatinny Arsenal, NJ. He is responsible for managing and directing advanced research programs in the areas of sensor based robotics, intelligent systems, digital controls and embedded decision support systems and has authored numerous publications in his areas of specialization. Prior to his current position Dr. Coleman was chief of the Weapon Control and Stabilization Group (1980-1986) and senior staff mathematician (1971-1980). He has been the recipient of three Army Research and Development Awards during this period. He was also Adjunct Assistant Professor of Applied Mathematics at the University of Iowa (1971-1976). From 1969-1971 he served as Captain in the U.S. Army Ordnance corps and received the Bronze Star during service in Vietnam.