Comparison of Standard and Modified Recursive State Estimation Techniques For Nonlinear Systems

Alexandros C. Charalampidis and George P. Papavassilopoulos

***Abstract*—This paper deals with recursive state estimation for nonlinear systems. A new set of σ-points for the Unscented Kalman Filter is proposed as well as a way to substitute a nonlinear output with a linear one. The importance of the function of the state which must be estimated is also illustrated and the need for taking it into account when designing the state estimator. All the proposed methods are compared with standard Extended Kalman Filter, Unscented Kalman Filter and Particle Filter with Sampling Importance Resampling using simulations. The results show that the modifications proposed in some cases lead to considerable reduction in estimation error.**

***Index Terms*—Kalman filtering, Nonlinear estimation, Nonlinear filters, State estimation.**

I. Introduction

ESTIMATING the state of a dynamical system is a common task. In practice, all measurements are noisy and all processes are affected by some kind of disturbance. For linear equations with normally distributed disturbance and measurement error stochastic processes it is well known that Kalman Filter [4] provides an exact solution of the problem. The problem can be also solved exactly if the state space is finite.

In this paper nonlinear discrete time systems with additive noise are considered. For these systems, even if all disturbances are normally distributed, the nonlinearities distort the distribution thus leading to non-normal distributions for the state of the system. For low dimensional systems it is feasible to approximate the exact state distribution by partioning the state space. Other approximate solutions, which do not suffer from the curse of dimensionality, are provided by the Extended Kalman Filter (EKF) [1], the Unscented Kalman Filter (UKF) [5] and various forms of Particle Filters (PF) [7].

For the UKF the selection of σ-points is an important issue. Several aspects of this selection are treated in [5]. A new selection algorithm is presented in this paper, and in a semi-realistic example, it outperforms the standard algorithm of [8], also used by [9] and most authors.

Apart from the system dynamics, nonlinearities may be present on the output equations. Under certain conditions, inverting the nonlinearity and assuming a linear output equation can help to avoid nonlinear overshoot-like phenomena. This is illustrated with an appropriate example, and it is shown using simulations that by this way the error of the EKF can be reduced significantly.

The authors are with the School of Electrical and Computer Engineering, National Technical University of Athens, Greece.
e-mail: alexchar@central.ntua.gr, yorgos@netmode.ntua.gr

The remainder of this paper is organized as follows. In Section II the problem formulation and basic results from probability theory are presented. The standard filtering techniques compared with the proposed ones are also presented. In Section III the proposed modifications are described. Section IV presents simulation examples used to compare the performance of the techniques under consideration. Conclusions are drawn in Section V.

II. Background

A. Problem Formulation

The dynamic systems considered are of the form

$$x_{k+1} = f(x_k) + w_k, \tag{1}$$

$$y_k = h(x_k) + v_k, \tag{2}$$

where x_k is the state of the system and y_k the measured output at the time system k. w_k is the disturbance, also referred to as process noise, and v_k is the measurement noise. In this paper only i.i.d. noise sequences will be considered. The initial condition x_0 follows a known distribution, and measurements are available from time $t = 1$ and onwards.

The problem is to estimate a function of the state, $z_k = g(x_k)$. The function g can be the identity, the output function h, or another function, although the last case is not addressed in this contribution. Since z_k will be a random variable, it may be desired to approximate its probability density function (pdf) or only some statistics of it, such as the expected value or the covariance. As it will be made apparent, the choice of g and of the statistics that must be estimated has heavy impact on the whole procedure.

Suppose now that $p_{X_0}(x_0)$ is the pdf of x_0, $p_V(v_k)$ is the pdf of the measurement noise and $p_W(w_k)$ is the pdf of the process noise. It holds $p_{Y|X}(y_k|x_k) = p_V(y_k - h(x_k))$ and $p_{X_{k+1}|X_k}(x_{k+1}|x_k) = p_W(x_{k+1} - f(x_k))$. The subscripts of probability density functions will be omitted for convenience. Let us define $y_{1:k} = \{y_1, y_2, \cdots, y_k\}$.

Then, using Bayes Rule [2], the following recursive equations hold:

$$p(x_{k+1}|y_{1:k}) = \int p(x_{k+1}|x_k)p(x_k|y_{1:k})dx_k, \tag{3}$$

$$p(x_{k+1}|y_{1:k+1}) = p(y_{k+1}|x_{k+1})p(x_{k+1}|y_{1:k})/c_k, \tag{4}$$

where

$$c_k = \int p(y_{k+1}|x_{k+1})p(x_{k+1}|y_{1:k})dx_k. \tag{5}$$

However the above integrals cannot be evaluated analytically. Numerical integration for a sufficiently dense mesh of x_k at each time step is also impractical, so (3)–(5) are mainly of theoritical interest.

B. Standard Filtering Techniques

In EKF the system equations (1)–(2) are linearized and then Kalman Filter (KF) Equations are used. Specifically, if $\hat{x}_k$ is the estimated mean of x_k after the correction step of time k and $\hat{x}^-_{k+1}$ is the estimated mean of x_{k+1} after the prediction step of time $k+1$, then for the prediction step Eq. (1) is linearized around $\hat{x}_k$ while for the correction step Eq. (2) is linearized around $y(\hat{x}^-_{k+1})$. Covariance matrices and correction gain are then calculated with KF Eqations. As is obvious, it is assumed that the derivatives of f and h are available.

The idea behind UKF is that it is easier to approximate a probability distribution than a nonlinear function. Thus system equations are not simplified but the prior distribution is approximated by a finite sum of Dirac deltas. The procedure which follows applies to systems with additive noise. For the general case, as well as for justification and criteria for parameter choice consult [5].

Suppose that after time k the covariance of the state is P_{x_k}. Let also n_x the dimension of the state space. Then $2n_x+1$ σ-points are calculated: $X_{0,k} = \hat{x}_k$, $X_{i,k} = \hat{x}_k + \gamma S_i, i = 1, \ldots, n_x$ and $X_{i,k} = \hat{x}_k - \gamma S_i, i = n_x+1, \ldots, 2n_x$, where S_i is the ith column of the matrix $S = \sqrt{P_k}$ and $\gamma = \sqrt{n_x+\lambda}, \lambda = \alpha^2(n_x+\kappa)-n_x$, α and κ being parameters.

For the prediction step, the σ-points are transformed through $x_{k+1} = f(x_k)$ and thus the covariance and mean before the correction can be calculated. Of course, the covariance of w_k is added to the covariance of the transformed σ-points. The weigths used for the mean and covariance calculations are

$$w_m^{(0)} = \frac{\lambda}{n_x+\lambda}, w_c^{(0)} = \frac{\lambda}{n_x+\lambda} + 1 - \alpha^2 + \beta \tag{6}$$

$$w_m^{(i)} = w_c^{(i)} = \frac{1}{2(n_x+\lambda)}, \tag{7}$$

where β is also a parameter.

For the correction step, with the new mean and covariance, $2n_x+1$ σ-points $X^-_{i,k+1}$, $i = 0, \ldots, 2n_x$ are analogously calculated. These are transformed through $y_{k+1} = h(x_{k+1})$ and thus the mean and covariance of the measurement vector can be calculated (for the covariance the covariance of v_k is added), as well as the cross covariance between the measurement and the state vector. Then the Kalman gain and the corrected estimates are computed according to KF Equations.

In this paper, UKF will be applied with the choice of parameters $\alpha = 1,\ \beta = 2,\ \kappa = 0$.

EKF and UKF are based on the assumption that the sequence of state distributions can be well approximated by a sequence of normal distributions. Since in many cases this assumption does not hold, and the form of the distribution is a priori unknown, there is need for a filtering technique which permits approximation of arbitrary distributions. In PF the state pdf is approximated by a number of particles each representing a Dirac delta with a corresponding weigth, i.e.

$$p(x_k|y_{1:k}) \approx \sum_{i=1}^{N} w_k^i \delta_{x_k^i}(x_k). \tag{8}$$

Then the mean value of a function g of the state can be estimated with the next equation.

$$\mathbb{E}[g(x_k)] \approx \sum_{i=1}^{N} w_k^i g(x_k^i) \tag{9}$$

It is reasonable to conjecture that with sufficiently many particles the approximation will be good, although asymptotic analysis of PF is a difficult problem [10].

There are many different algorithms [7] for the update of w_k^i and x_t^i. For the comparison purposes of this paper only Sampling Importance Resampling (SIR) Filtering, which consists the first such PF and was proposed by [6], will be used. The resampling step is implemented with stratified resampling [11]. Since this paper does not propose any novelty on PF, no further details are presented here.

III. Proposed Modifications

A. Inverting the Output Equation

Assume that EKF is used to estimate the state of the system described by Eqs. (21)–(22). (22) has the form of (2) with $h(x) = x^3$. Suppose that for some time instant k it holds $\hat{x}^-_k = 1, P^-_{x_k} = 1, x_k = 2$ and $y_k = 8$, while the variance of v_k is equal to 1. Then since $h'(x) = 3x^2$, $h'(\hat{x}^-_k) = 3$ and the Kalman gain will be equal to $\frac{3}{3^2+1}$. Since the predicted output was equal to 1, EKF yields $\hat{x}_k = 1 + 0.3(8-1) = 3.1$.

The posterior value for the output is greater than the observed, yet this is not due to measurement error and the predicted value was less than the observed. This happens because the first order approximation of h obtained with its derivative is valid only locally. Thus because of the nonlinearity of the output function, the Filter although supposed to smooth the observed data, suffers from this overshoot-like phenomenon.

A remedy of this problem is inverting the output equation. If there was no measurement error, one would set $x_k = \sqrt[3]{y_k}$, where the cubic root of a negative number is defined appropriately. Now it holds $x_k = \sqrt[3]{y_k + v_k}$. Since v_k is a random variable, $\sqrt[3]{y_k + v_k}$ is also a random variable. It is not computationally demanding to calculate its mean and variance. For this purpose, a suitable set $\{V_1, V_2, \cdots, V_N\}$ is determined and then for each step from the corresponding values $\{\sqrt[3]{y_k + V_1}, \cdots, \sqrt[3]{y_k + V_N}\}$ the mean s_k and variance l_k are extracted. Explain how $\{V_1, V_2, \cdots, V_N\}$ is determined?

EKF is then applied to the system whose dynamics is given by (21) and its output equation is

$$s_k = x_k + r_k, \tag{10}$$

where the variance of r_k is equal to l_k, which is obtained, together with s_k, from the inversion step.

The effectiveness of the proposed technique is illustrated in Fig. 1. It shows that after the inversion of the output equation the overshoot phenomenon is no more present. The idea of output inversion can be also applied to UKF and Fig. 1 shows that it is effective for UKF, too. In another realization of the experiment, presented in Fig. 2 EKF is trapped to zero while UKF suffers from overshoot. Again, application of output inversion is effective for both EKF and UKF. Numerical values of the prediction error for all filtering techniques tested

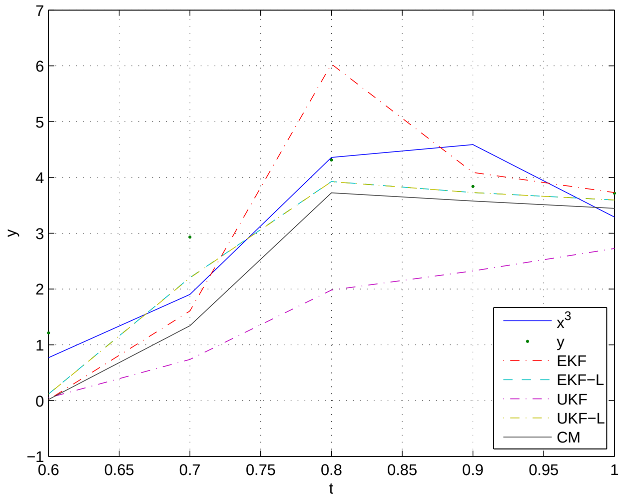

Fig. 1. Time series of output data and various estimation methods illustrating the overshoot-like phenomenon for EKF

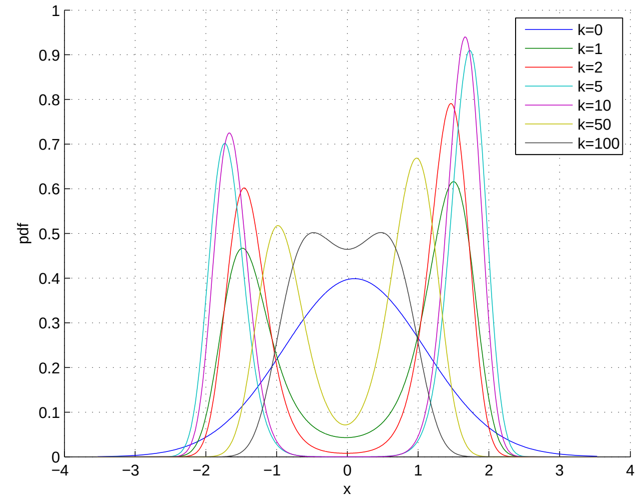

Fig. 3. Probability density function for various k in a random run

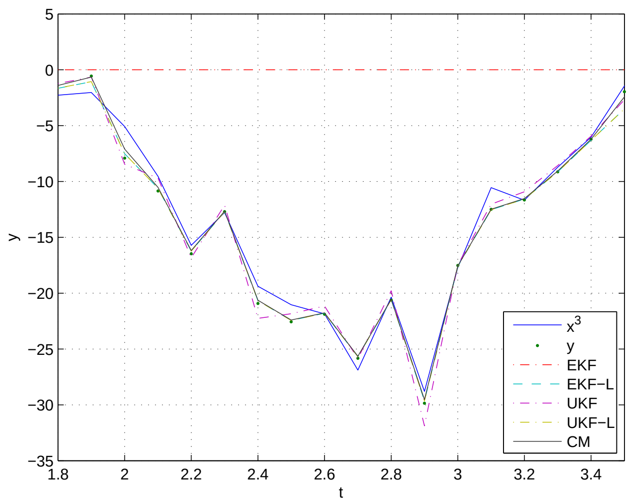

Fig. 2. Time series of output data and various estimation methods illustrating the overshoot-like phenomenon for UKF and the possibility of EKF being trapped to zero

are presented in section IV. Transfer here the inversion of subsection IV-C?

B. Estimating a Specific Function of the State

Let us consider the system of (21) and (23). Suppose that w_k and v_k are i.i.d. gaussian processes with variance Q and R respectively. (23) has the form of (2) with $h(x) = x^2$. Suppose also that it is known that $x_0 = 0$. Then $\hat{x}_1^- = 0$ and thus $h'(\hat{x}_1^-) = 0$ yielding zero Kalman gain and $\hat{x}_1 = 0$. Recursively one obtains $\hat{x}_k = 0$, whatever the output sequence is.

On the other hand, for every k the pdf of the state is even. In fact, if the ouput sequence is $\{y_1, y_2, \cdots, y_k\}$, then for every pair of sequences $\{x_1, x_2, \cdots, x_k\}$ and $\{v_1, v_2, \cdots, v_k\}$ consistent with $\{y_1, y_2, \cdots, y_k\}$, the pair $\{-x_1, -x_2, \cdots, -x_k\}$ and $\{v_1, v_2, \cdots, v_k\}$ is also consistent with the output sequence and equally probable. Marginalizing over possible $\{x_1, x_2, \cdots, x_{k-1}\}$ and $\{v_1, \cdots, v_{k-1}\}$ asserts that $p(x_k) = p(-x_k)$. Thus the expected value of x_k is 0, exactly as predicted by EKF or UKF. However, the pdf, which for $k = 0,\ 1,\ 2,\ 5,\ 10,\ 50,\ 100$ (0 corresponds to the prior) in a random run is presented at Fig. 3 is even for $k = 1$ not at all close to the pdf of a normal distribution. Thus, if it is not the mean of x_k that must be estimated but another quantity, such as the mean of x_k^2, EKF and UKF perform poorly. It is noted that even if $x_0 \neq 0$, since for some k the value of x_k will be close to 0, the same problem will arise.

An alternative way to estimate x_k^2 follows. Set $z_k = x_k^2$. Then if $|z_k| < \frac{1}{3\alpha\delta}$ it holds

$$\begin{aligned} z_{k+1} &= (x_k - \alpha\delta x_k^3 + w_k)^2 = \\ &= x_k^2 - 2\alpha\delta x_k^4 + \alpha^2\delta^2 x_k^6 + w_k^2 + 2(x_k - \alpha\delta x_k^3)w_k = \\ &= z_k - 2\alpha\delta z_k^2 + \alpha^2\delta^2 z_k^3 + w_k^2 + 2(x_k - \alpha\delta x_k^3)w_k, \end{aligned} \tag{11}$$

while in different case it holds

$$z_{k+1} = x_\infty^2 - 2\alpha\delta x_\infty^4 + \alpha^2\delta^2 x_\infty^6 + w_k^2 \pm 2(x_\infty - \alpha\delta x_\infty^3)w_k, \tag{12}$$

where x_∞ is defined in (20). The expected value of the unknown term is $\mathbb{E}[w_k^2 + 2(x_k - \alpha\delta x_k^3)w_k] = Q$. As for its variance, it depends on $|x_k|$. It is possible to create off-line a look-up table and use it to find the variance at each time step. More details about the look-up table? Thus an estimate of x_k^2 is obtained by estimating the state of the system defined by (13)–(14)

$$z_{k+1} = z_k - 2\alpha\delta z_k^2 + \alpha^2\delta^2 z_k^3 + Q + n_k \tag{13}$$

$$y_k = z_k + v_k, \tag{14}$$

where n_k is zero-mean but its distribution and variance depend on z_k.

For numerical results and comparison of all filtering techniques tested in this paper for this problem see subsection IV-B.

C. A New σ-Point Set

1) One-Dimensional Case: The algorithm of subsection II-B approximates the normal distribution with mean M and variance P with three σ-points, at M, $M-\gamma\sqrt{P}$ and $M+\gamma\sqrt{P}$ where γ as well as the corresponding weights are given in the same subsection.

More points could be used to obtain a better approximation. Let N the desired number of σ-points. The following algorithm provides a reasonable way to approximate the normal distribution with mean M and variance P using N points. Restate the last sentence? Let $\{p_i = \frac{i}{N+1}, i = 1 \dots N\}$. Then calculate $\{y_i = erf^{-1}(p_i), i = 1 \dots N\}$. Then the set $\{y_i, i = 1 \dots N\}$ is zero-mean but not necessarily of variance 1. A set of variance 1 is obtained by

$$\{x_i = \sqrt{\frac{N}{\sum_{i=j}^{N} y_j^2}} y_i, i = 1 \dots N\}. \quad (15)$$

The proposed σ-point set is then defined by

$$\{\sigma_i = M + \sqrt{P} x_i, i = 1 \dots N\}, \quad (16)$$

and its mean and variance are equal to M and P respectively. It is noted that both for the one-dimensinal and the multi-dimensional case, the proposed point have equal weights.

2) Multi-Dimensional Case: Let M the vector mean and P the covariance matrix of a n-dimensional normal distribution. Since P is symmetric it is possible to find n orthonormal eigenvectors $\{v_i, i = 1 \dots n\}$ with corresponding eigenvalues $\{\lambda_i, i = 1 \dots n\}$. P is positive definite therefore all eigenvalues are positive. Let $\{z_i, i = 1 \dots n\}$ n i.i.d. random variables following the standard normal distribution. Then the random vector $M + \sum_{i=1}^{n} z_i v_i$ is normally distributed with mean M and covariance P.

Since the standard normal distribution is approximated by the set defined by (15), the following set, consisted of N^n members, can be considered as an approximation of the n-dimensional normal distribution with mean M and covariance P.

$$\{\sigma_{i_1, i_2, \dots, i_n} = M + \sum_{j=1}^{n} x_{i_j} v_i, i_j = 1 \dots N, j = 1 \dots n\} \quad (17)$$

Its mean and covariance are indeed M and P respectively.

For large n the cardinal of the set, N^n, will be much greater than $2n+1$ even for $N = 2$. Apart from that, the above algorithm includes computation of the eigenvalues of P while for the set presented in subsection II-B it is only needed to compute the square root of a positive definite symmetric matrix, a task that can be accomplished in $O(n^3)$ time using Cholesky Factorization. However, numerical experiments with the four-dimensional example of subsection IV-C show that for that case the extra computational cost associated with the proposed σ-point set is not prohibitive.

The following example shows that the use of all combinations of i_j in $\sum_{j=1}^{n} x_{i_j} v_i$ can help compute the covariance of the output much more accurately than with the σ-point set of subsection II-B. Suppose that $(x_1, x_2)^T$ is normally distributed with zero mean and covariance matrix equal to the identity matrix, and that the mean and variance of $f(x_1, x_2) = x_1 x_2$ must be calculated. The standard algorithm uses five points, $\{(0,0), (0,1), (0,-1), (1,0), (-1,0)\}$. All of them are mapped by f to 0 and thus the variance of f is estimated to be 0. The proposed algorithm with $N = 2$ yields four points, $\{(-1,1), (1,1), (1,-1), (-1,-1)\}$. Two of them are mapped to 1 and two to -1, thus the variance is estimated to be 1, which is the exact value. Both algorithms correctly estimate the mean to be 0.

Numerical results of the use of the proposed σ-point set in recursive state estimation are provided in the next section.

TABLE I
TABLE OF RMS ESTIMATION ERROR AND COMPUTATION TIME FOR THE EXAMPLE OF SUBSECTION IV-A, CASE I

Estim. Technique	x Est. Error	y Est. Error	Comp. Time (ms)
EKF	0.8440	12.96	0.66
UKF	0.3939	1.3300	1.74
MUKF	0.3937	1.3256	1.58
EKF-L	0.4272	0.7881	2.55
UKF-L	0.4266	0.7872	3.14
PF/SIR(50)	0.3936	0.8358/0.8572	0.94
PF/SIR(1000)	0.3814	0.7266/0.7497	5.18
CM(50)	0.3819	0.9123/0.9329	0.23
CM(1000)	0.3804	0.7224/0.7455	11.56

IV. SIMULATION EXAMPLES

A. One-Dimensional Example with Invertible Ouput Equation

Consider the system dynamics described by (18), where w is the disturbance.

$$\dot{x} = -\alpha x^3 + w \quad (18)$$

The simplest discrete time approximation with time step δ is given by the following equation.

$$x_{k+1} = x_k - \alpha \delta x_k^3 + w_k \quad (19)$$

However $x - \alpha \delta x^3$ is not a monotonous function of x, although it should be so as to yield an acceptable approximation of the dynamics of (18). Since it has a maximum for

$$x_\infty = \frac{1}{\sqrt{3\alpha\delta}} \quad (20)$$

the following approximation captures better the continuous time dynamics.

$$x_{k+1} = f(x_k) + w_k, f(x) = \begin{cases} x - \alpha\delta x^3 & |x| < x_\infty, \\ -\frac{2}{3} x_\infty & x \le -x_\infty, \\ \frac{2}{3} x_\infty & x \ge x_\infty. \end{cases} \quad (21)$$

Let us consider therefore the problem of estimating the state of the system with dynamics given by (21) and output given by (22)

$$y_k = x_k^3 + v_k. \quad (22)$$

The system has been simulated under the following conditions. The parameters are set to $\alpha = 0.1$, $\delta = 0.1$. The initial condition x_0 is normally distributed with mean 0.1 and variance 1, and independent of w_k and v_k, who are also

TABLE II
TABLE OF RMS ESTIMATION ERROR AND COMPUTATION TIME FOR THE EXAMPLE OF SUBSECTION IV-A, CASE II

Estim. Technique	x Est. Error	y Est. Error	Comp. Time (ms)
EKF	0.4538	1.2696	0.67
UKF	0.3459	1.5119	1.71
MUKF	0.3455	1.4845	1.52
EKF-L	0.3864	0.4056	2.45
UKF-L	0.3861	0.4042	3.07
PF/SIR(50)	0.3449	0.3888/0.3885	0.95
PF/SIR(1000)	0.3272	0.3477/0.3508	5.37
CM(50)	0.3268	0.3492/0.3523	0.22
CM(1000)	0.3267	0.3467/0.3498	11.23

assumed to be normal and independent of each other. They are i.i.d. and both zero-mean. The variance of v_k is assumed equal to 1, while for the first case studied, the variance of w_k is equal to $\sqrt{0.1}$. The system is simulated up to time $t = 10$.

In order to compare the different estimation methods, the experiment has been run 100 times. The simulated mean values of x^2 and y^2 are 1.53 and 26.78 respectively. Table I presents the rms values of the estimation error for x and y, as well as the mean computation time per step in msec. All computation times in this paper have been recorded for a 32-bit PC clocked at 1.8GHz and running MATLAB 7.2 for Linux.

EKF and UKF are standard Extended and Unscented Kalman Filter respectively. MUKF stands for Modified UKF. The modification implied is the use of the proposed σ-point set with $N = 5$ in the prediction step instead of the standard set. EKF-L and UKF-L are EKF and UKF with inverted output equation, i.e. with output equation (10) instead of (22). PF/SIR stands for Particle Filter with Sampling Importance Resampling, while the number in parenthesis denotes the number of particles. Since PF provides an approximation of the distribution, y can be estimated using (9). The number at the left of the slash is $\mathbb{E}[h(x_k)]$ while the number at the right of the slash is $h(\mathbb{E}[(x_k)])$ and, as expected, the former is smaller than the latter.

Finally, the state space has been partioned using a constant mesh (CM). The number in parenthesis is the number of intervals used. The dynamics are then described by a Markov matrix, and the prediction step can be accomplished as a multiplication of a matrix by a vector. Then, for the correction step, the probability of each interval is multiplied by the likelihood of the observed value with respect to its center, and then the probabilites are normalized. The numbers at the left and right of the slash have the same meaning as for PF, and the same comment applies to this case, too.

The results show that EKF cannot provide a satisfactory estimate for x or y. This is due to the nonlinear effect described in subsection III-A. UKF is also affected by the same phenomenon, and the mean error for y is still high, although x is estimated well. MUKF is slightly better than UKF, but still unsatisfactory for y. Inverting the output equation reduces substantially the estimation error for y for both EKF and UKF, but increases the error for x. PF and CM, with many particles and intervals respectively, perform better than UKF-L, but the fact that the y-error for EKF-L and UKF-L is smaller than that of PF(50) and CM(50) shows the strength of the proposed technique. It is also interesting that although CM(50) yields x-error comparable to that of PF(1000), its y-error is higher than that of PF(50).

The system has been also studied for variance of w_k equal to 0.1. For this case the simulated mean values of x^2 and y^2 are 0.54 and 2.67 respectively, while the results are presented in Table II. Again, EKF performs poorly. UKF now estimates x well, but its y-error is greater even than that of EKF. MUKF yields only slight error decrease. Inverting the output equation permits good estimation of y, but deteriorates the estimation of x. Finally, it can be observed that for this case CM estimates y better than PF.

B. One-Dimensional Example with Non-Invertible Ouput Equation

In this subsection the system with dynamics given by (21) and output given by (23) is studied.

$$y_k = x_k^2 + v_k \tag{23}$$

The parameters as well as the distribution of x_0 are that of the preceding example. The distribution of w_k and v_k is that of the first case of that example. The system is simulated up to time $t = 10$ and 100 runs have been made, as for the system of the previous example.

The simulated mean values of x^2 and y^2 are 1.46 and 6.00 respectively. The results are presented in Table III. Only the estimation error for y is given, for the reasons explained in Subsection IV-B. The values on the second and third column correspond to the methods applied to (13)–(14) (real system), while those of the fourth and fifth column correspond to (21) and (23) (virtual system). When estimating the state of the virtual system, for simplification n_k is assumed normally distributed. The computation times given are per step and in msec.

EKF behaves well for both systems, while UKF only for the virtual one. MUKF leads to considerable error decrease in comparison with UKF only for the real system. The fact that the estimation error using EKF, UKF and MUKF for the virtual system is close to that of CM or PF for the real one illustrates the efficacy of the proposed technique. It is also interesting that for the virtual system PF and CM perform poorly. Finally, it is noted that the computation time for CM in the case of the virtual system is high because the stochastic matrix depends on the process noise variance, and thus must be calculated at each step.

C. Four-Dimensional Semi-Realistic Example

In the last subsection a semi-realistic example of an armature controlled DC motor with a sin/cos encoder is studied. The field current is assumed constant and magnetic nonlinearities are neglected. θ denotes angular position, ω angular speed and i armature current. e is the voltage applied to the armature, used to control the motor, and for this example it will be assumed constant. The damping coefficient $b(t)$ is supposed to vary periodically, although the corresponding dynamics is

TABLE III
TABLE OF RMS ESTIMATION ERROR AND COMPUTATION TIME FOR THE EXAMPLE OF SUBSECTION IV-B

Estim. Techn.	(21), (23)	Comp. Time	(13)–(14)	Comp. Time
EKF	0.6840	0.60	0.6466	1.38
UKF	1.7130	1.54	0.6468	2.39
MUKF	1.2121	1.40	0.6417	2.20
PF/SIR(30)	0.6199	0.81	0.6803	1.49
PF/SIR(100)	0.6039	2.47	0.6623	1.77
CM(30)	0.6083	0.19	0.8031	15.51
CM(100)	0.6035	2.74	0.7694	41.94

TABLE IV
TABLE OF RMS ESTIMATION ERROR AND COMPUTATION TIME FOR THE EXAMPLE OF SUBSECTION IV-C

Estim. Technique	θ	ω	i	ω_b	Comp. Time
EKF	0.0569	0.0344	0.0017	0.1562	3.53
UKF	0.0067	0.0230	0.0071	0.1318	13.60
MUKF	0.0026	0.0172	0.0012	0.0881	20.61
EKF-L	0.0151	0.0260	0.0008	0.1322	2.86
UKF-L	0.0069	0.0231	0.0072	0.1318	12.85
PF/SIR(100)	0.0068	0.0596	0.0128	0.1980	5.59
PF/SIR(1000)	0.0067	0.0287	0.0040	0.1387	28.10

subject to disturbance. Then the system dynamics are given by (24)–(28).

$$\dot{\theta} = \omega \tag{24}$$
$$J\dot{\omega} = -b(t)\omega + Ki \tag{25}$$
$$L\dot{i} = e - K\omega - Ri \tag{26}$$
$$\dot{\phi}_b = \omega_{b0} + w \tag{27}$$
$$b(t) = b_0 + b_1 cos(\phi_b) \tag{28}$$

The only quantity measured is the angular position, using a sin/cos encoder. It is assumed that measurements are taken every $\delta = 0.01 sec$ according to (29)–(30), where $v_{1,k}$ and $v_{2,k}$ are i.i.d. gaussian random variables with zero mean and variance R_v.

$$y_{1,k} = sin(\theta_k) + v_{1,k} \tag{29}$$
$$y_{2,k} = cos(\theta_k) + v_{2,k} \tag{30}$$

In discrete time the dynamics are approximated by the following equations.

$$\theta_{k+1} = \theta_k + \delta\omega \tag{31}$$
$$\omega_{k+1} = \omega_k + \delta(-b_k\omega + Ki)/J \tag{32}$$
$$i_{k+1} = i_k + \delta(e - K\omega - Ri)/L \tag{33}$$
$$\phi_{b,k+1} = \phi_{b,k} + \delta\omega_{b0} + w_k \tag{34}$$
$$b_k = b_0 + b_1 cos(\phi_{b,k}) \tag{35}$$

Thus the problem is to estimate the state of the system with state equations (31)–(34), where b_k is defined by (35), and output equations (29)–(30).

As in subsection III-A, Eqs. (29)–(30) can be inverted to provide a virtual measurement of θ, but in this case it is not so straightforward. It holds $\theta = arctan(y_1 - v_1, y_2 - v_2)$, where by $arctan(y, x)$ is denoted the angle whose sine and cosine are equal to $\frac{y}{\sqrt{x^2+y^2}}$ and $\frac{y}{\sqrt{x^2+y^2}}$ respectively. Let $R = \sqrt{y_1^2 + y_2^2}, \phi = arctan(y_1, y_2)$ so that $y_1 = Rsin\phi$ and $y_2 = Rcos\phi$. Define the random variables $n = cos\phi v_1 - sin\phi v_2$ and $r = -sin\phi v_1 - cos\phi v_2$. They are normally distributed and it can be verified through simple calculations that their variance is equal to R_v as well as that they are uncorrelated and thus independent.

Let us define the random variable $q = -\frac{n}{r+R}$. Then

$$\begin{aligned}\theta &= arctan(Rsin\phi + rsin\phi - ncos\phi, Rcos\phi + rcos\phi + nsin\phi)\\ &= arctan\big((R+r)sin\phi - ncos\phi, (R+r)cos\phi + nsin\phi\big)\\ &= arctan(sin\phi + qcos\phi, cos\phi - qsin\phi).\end{aligned} \tag{36}$$

Using the trigonometric identity

$$tan(a+b) = \frac{tan(a) + tan(b)}{1 - tan(a)tan(b)} \tag{37}$$

the reader can check that

$$\theta = arctan(sin\phi + qcos\phi, cos\phi - qsin\phi) = \phi + arctan(q). \tag{38}$$

Since $arctan$ is an odd function and the pdf of n is also odd it follows that $\mathbb{E}[arctan(q)] = 0$ thus $\mathbb{E}[\theta] = \phi$. The variance of $arctan(q)$ depends on R.

It must be noted that ϕ is also the maximum likelihood estimator of θ. Indeed, the likelihood function is

$$\frac{1}{\sqrt{2\pi R_v}} e^{-\frac{(y_1 - sin\phi)^2}{2R_v}} \frac{1}{\sqrt{2\pi R_v}} e^{-\frac{(y_2 - cos\phi)^2}{2R_v}}. \tag{39}$$

Thus it is maximized when the expression $(y_1 - sin\phi)^2 + (y_2 - cos\phi)^2$ is minimized. But

$$\begin{aligned}&(y_1 - sin\phi)^2 + (y_2 - cos\phi)^2 =\\ &= y_1^2 + y_2^2 + sin^2\phi + cos^2\phi - 2(y_1 sin\phi + y_2 cos\phi) =\\ &= y_1^2 + y_2^2 + 1 - 2(y_1 sin\phi + y_2 cos\phi).\end{aligned} \tag{40}$$

It is left to the reader to verify that $(y_1 sin\phi + y_2 cos\phi)$ is maximized for $\phi = arctan(y_1, y_2)$ and thus complete the argument.

A final remark is that in order to avoid 2π jumps, at each time step an integer multiple of 2π is added to $arctan(y_1, y_2)$ so as to make the sum as close as possible to the previous estimated value of θ. Explain Further?

The system parameters for the simulations performed have been selected as follows (in SI units). $K = 0.035$, $L = 0.5$, $R = 2$, $J = 0.03$, $b_0 = 0.01$, $b_1 = 0.003$, $\omega_{b0} = 10$, $e = 20$. The initial condition follows the normal distribution with zero mean and covariance matrix $0.01I_4$. The variance of w_k, $v_{1,k}$ and $v_{2,k}$ is equal to 10^{-4}. All noise sequences are normally distributed, i.i.d. and independent of each other.

In Fig. 4 waveforms of the estimation error for the angular velocity and for all techniques applied to this problem are

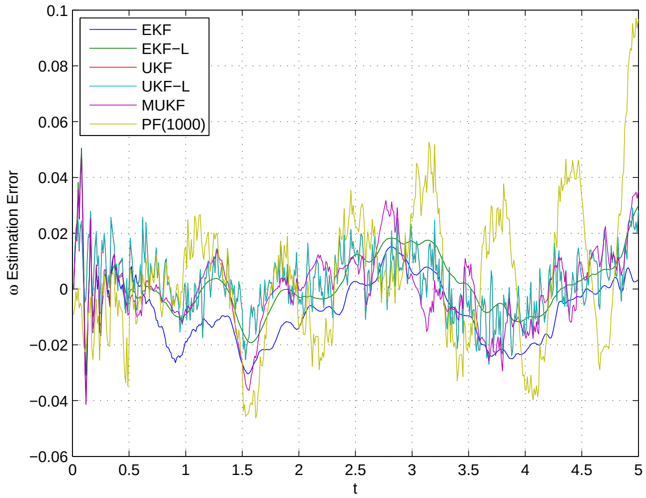

Fig. 4. ω Estimation Error for various methods in a random run

shown. It is clear that MUKF performs better than any other filter. The results from 100 runs are presented in Table IV. The values of the estimation error given are rms values, while the rms values for the four variables are equal to 38.3, 18.2, 9.4 and 28.9.

For MUKF N was chosen equal to 3. Partioning the state space with a constant mesh was not a choice for this case, because the problem now is four-dimensional. MUKF gives by far the best estimate for θ, the best estimate for i and ω_b, while for ω it is bettered only by EKF-L. Using a suitable virtual output as described in this subsection decreases considerably the error for EKF. UKF instead, which handles better nonlinearities by its own, does not benefit from the inversion. At this point it must be noted that for simplification the variance of q in (38), which depends on R, has been treated as constant, with value that for $R = 1$.

The computation times given are in msec. MUKF is more demanding than UKF, but for this case both times are of the same order. It is interesting that PF/SIR(1000) although the most computationally costly method does not give the best estimate for any variable in this example.

The results of this subsection strongly encourage the use of the proposed methods.

V. Conclusion

This paper shows that inverting the output equation to provide a virtual linear output can diminish nonlinear phenomena and decrease output prediction error. Using the dynamics of a specific function of the state instead of those of the state for estimation purposes has been also illustrated through an example, for which it leads to better performance of EKF and UKF. Finally, a new σ-point set for UKF is proposed, and shown to outperform the standard set in a semi-realistic example.

Further research could be studying the general case for output equation inversion and estimating a specific function of the state. Since the cardinal of the proposed σ-point set increases exponentially with the dimension, it would be interesting to exploit sparsity patterns in the Jacobian of f in (1) to permit its use in high-order systems.

References

[1] B. D. O. Anderson and J. B. Moore, *Optimal Filtering.* Englewood Cliffs, NJ: Prentice-Hall, 1979.

[2] P. Billingsley, *Probability and Measure, 3rd Ed.* Hoboken, NJ: Wiley-Interscience, 1995.

[3] R. v. .d. Merwe, A. Doucet, N. d. Freitas and E. Wan, "The Unscented Particle Filter," Tech. Rep. CUED/F-INFENF/TR 380, Cambridge University Engineering Department, Aug. 2000.

[4] R. E. Kalman, "A new approach to linear filtering and prediction problem," in *ASME Journal of Basic Engineering*, Vol. 82, Issue 1, pp. 35-45, Mar. 1960.

[5] S. Julier and J.K. Uhlmann, "Unscented filtering and nonlinear estimation," in *Proceedings of the IEEE*, Vol. 92, Issue 3, pp. 401-422, Mar. 2004.

[6] N. J. Gordon, D. J. Salmond and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian state estimation," in *IEE Proceedings-F*, Vol. 140, Issue 2, pp. 107-113, Apr. 1993.

[7] M. Sanjeev Arulampalam, S. Maskel, N. Gordon and T. Clapp, "A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking," in *IEEE Transactions on Signal Processing*, Vol. 50, Issue 2, pp. 174-188, Feb. 2002.

[8] S. Julier, J. Uhlmann and H.F. Durrant-Whyte, "A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators," in *IEEE Transactions on Automatic Control*, Vol. 45, Issue 3, pp. 477-482, Mar. 2000.

[9] R. Kandepu, L. Imsland, and B. A. Foss, "Constrained State Estimation Using the Unscented Kalman Filter," in *Proc. 16th Mediterranean Conference on Control and Automation,*, pp. 1453-1458, 2008.

[10] X.-L. Hu, T. B. Schon, and L. Ljung, "A Basic Convergence Result for Particle Filtering," in *IEEE Transactions on Signal Processing*, Vol. 56, Issue 4, pp. 1337-1348, Apr. 2008.

[11] G. Kitagawa, "Monte Carlo Filter and Smoother for Non-Gaussian Nonlinear State Space Modules," in *Journal of Computational and Graphical Statistics*, Vol. 5, Issue 1, pp. 1-25, Mar. 1996.