

Mean-Square Estimation of Nonlinear Functionals via Kalman Filtering

Won Choi ¹, Vladimir Shin ² and Il Young Song ^{3,*}

¹ Department of Mathematics, Incheon National University, 119 Academy-ro, Yeonsu-gu, Incheon 406-772, Korea; choiwon@inu.ac.kr

² Department of Information and Statistics, Research Institute of Natural Science, Gyeongsang National University, 501 Jinjudaero, Jinju, Gyeongsangnam-do 660-701, Korea; vishin@gnu.ac.kr

³ Laser & Sensor System Team, Hanwha Daejeon Plant, 305 Pangyo-ro, Bundang-gu, Seongnam-si 13488, Korea

* Correspondence: com21dud@hanwha.com; Tel.: +82-10-2696-1875

Received: 3 October 2018; Accepted: 6 November 2018; Published: 12 November 2018

Abstract: This paper focuses on estimation of a nonlinear functional of state vector (NFS) in discrete-time linear stochastic systems. The NFS represents a nonlinear multivariate functional of state variables, which can indicate useful information of a target system for control. The optimal mean-square estimator of a general NFS represents a function of the Kalman estimate and its error covariance. The polynomial functional of state vector is studied in detail. In this case an optimal estimation algorithm has a closed-form computational procedure. The novel mean-square quadratic estimator is derived. For a general NFS we propose to use the unscented transformation to calculate an optimal estimate. The obtained results are demonstrated on theoretical and practical examples with different types of NFS. Comparative analysis with suboptimal estimators for NFS is presented. The subsequent application of the proposed estimators to linear discrete-time systems demonstrates their practical effectiveness.

Keywords: nonlinear functional; quadratic functional; mean square error; discrete-time system; Kalman filtering; multivariate normal distribution

1. Introduction

This paper is focused on the development of the optimal mean-square estimator for a nonlinear functional of state vector (NFS) in linear discrete-time stochastic systems. In this section, background material, a concise literature review, the formulation of the problem of interest for this study, the scope and the contributions of this investigation, and the organization of the paper are provided.

1.1. Background and Significance

Kalman filtering and its variations are well-known state estimation techniques in wide use in a variety of applications, such as navigation, target tracking, unmanned vehicle state estimation, communications engineering, air traffic control, biomedical and chemical processing, and many others [1–10].

However, some applications require the estimation of not only a state vector $x_t \in \mathbb{R}^n$ but also NFS, $z_t = f(x_t)$, which expresses practical and worthwhile information for control systems. For example, the Euclidean norm, $f(x_t) = \|x_t - x_t^n\|$, represents distance between the current x_t and nominal x_t^n states, respectively.

1.2. Literature Review

The problem of estimation of nonlinear functionals with unknown parameters and signals based on the minimax theory has been studied by many authors [11–14] and the references therein. Estimation of parameters of nonlinear functional model with known error covariance matrix is presented in [15]. Minimax quadratic estimate for integrated squared derivative of a periodic function is derived in [16]. In [17,18], the optimal matrix of a quadratic functional is determined based on cumulant criteria. Estimation of penalty considered as a quadratic cost functional for a quantum harmonic oscillator is given in [19]. Estimators for integrals of nonlinear functions of a probability density are developed in [20,21]. We also mention estimation of nonlinear functionals of spectral density, periodogram of stationary linear signals [22–24]. Some extension of these results is obtained by [25]. In [26] an unknown distance between a target and radar is approximated by Taylor polynomial to subsequent estimation of its coefficients. For algorithms and theory of estimation information measures representing a nonlinear functional of signals the reader is referred to [27,28]. However, most authors have not focused on estimation of nonlinear functionals of vector signals in dynamical models defined by stochastic equations. The first step in this direction is given in [29]. However, main results in [29] are obtained by estimation of an unknown signal with time delays using the robust receding horizon approach, rather than estimation of a nonlinear functional. The current paper focuses on the development of an analytical and computational framework for the optimal mean square estimation of a nonlinear functional, and the original signal is estimated by using the Kalman filter.

1.3. Formulation of the Problem of Interest for this Investigation

Consider a discrete-time linear stochastic system

$$\begin{aligned} x_{k+1} &= F_k x_k + G_k v_k, \quad k = 0, 1, \dots, \\ y_k &= H_k x_k + w_k, \end{aligned} \quad (1)$$

where $x_k \in \mathbb{R}^n$ is a state vector, $y_k \in \mathbb{R}^m$ is a measurement vector, the system noise $v_k \in \mathbb{R}^q$ and the measurement noise $w_k \in \mathbb{R}^m$ are uncorrelated white Gaussian noises with zero mean and covariances $Q_k \in \mathbb{R}^{q \times q}$ and $R_k \in \mathbb{R}^{m \times m}$, respectively, and F_k , G_k , and H_k are matrices with compatible dimensions. The initial state, $x_0 \sim N(m_0, C_0)$, is assumed to be Gaussian and uncorrelated with v_k and w_k , and $m_0 = E(x_0)$, $C_0 = \text{Cov}(x_0, x_0)$.

A problem associated with the system portrayed in Equation (1) is that of estimation of the NFS:

$$z_k = f(x_k): \mathbb{R}^n \rightarrow \mathbb{R} \quad (2)$$

from the overall noisy sensor measurements $Y^k = \{y_1, \dots, y_k\}$.

1.4. Scope and Contributions of this Investigation

The aim of this paper is to develop an optimal two-stage minimum mean square error (MMSE) estimator for an arbitrary NFS in a linear stochastic system, and investigate special polynomial and quadratic estimators for which one can obtain important mean square estimation results. The main contributions of the paper are the following:

- (1) This paper studies the MMSE estimation problem of an NFS within the discrete-time Kalman filtering framework. Using the mean square estimation approach, an optimal two-stage nonlinear estimator is proposed.
- (2) The optimal MMSE estimator for polynomial functionals (such as quadratic, cubic and quartic) is derived. We establish that the polynomial estimator representing a compact closed-form formula depends only on the Kalman estimate and its error covariance.
- (3) An important class of quadratic estimators is comprehensively investigated, including derivation of novel matrix equations for the estimated and true mean square error.
- (4) Performance of the proposed estimators for real NFS illustrates their theoretical and practical usefulness.

1.5. Organization of the Paper

This paper is organized as follows. Section 2 presents two motivating examples. The main theoretical results are presented in Section 3, which is divided into four subsections: (i) main idea and general solution; (ii) optimal closed-form MMSE estimator for quadratic functional; (iii) optimal closed-form MMSE estimator for polynomial functional; and, (iv) application of the unscented transformation for estimation of general NFS. Simulation analysis and verification of the proposed estimators are given in Section 4, which is divided into four subsections: (i) estimation of distance between unknown and moving points; (ii) estimation of power and modulus of unknown signal; (iii) estimation of total kinetic energy of wind tunnel system; and, (iv) estimation of size of oil spread contour. In Section 5, the summary of this investigation, the conclusions drawn in this paper, and the future direction of research are provided.

2. Motivating Examples

Two examples illustrate an NFS and how it may be used in practice. The first motivating example of the NFS is kinematic quantities of projectile motion. State vector of an object $x \in \mathbb{R}^4$, an angle (θ) and range (r) are shown in Figure 1.

$$\begin{aligned} f(x) &= \theta = \tan^{-1}(x_2/x_1), \\ f(x) &= r = \sqrt{x_1^2 + x_2^2}, \\ f(x) &= |\vec{v}| = \sqrt{x_3^2 + x_4^2}, \quad x_3 = v_1, \quad x_4 = v_2. \end{aligned} \quad (3)$$

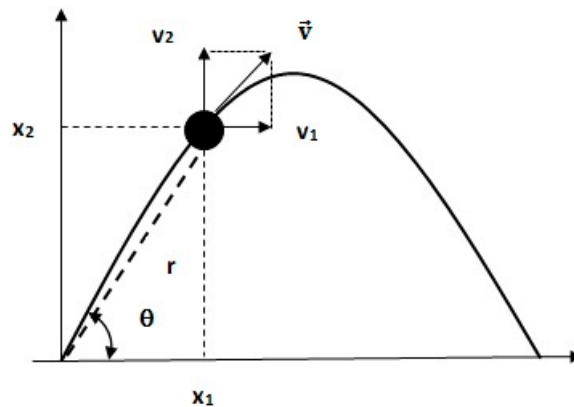


Figure 1. Kinematic quantities of projectile motion.

The second example can be an arbitrary quadratic form, $f(x_t) = x_t^T \Omega x_t$, representing an energy-like function of an object or the Euclidean square distance, $f(x_t) = d^2(x_t, x_t^n)$, between the current x_t and nominal x_t^n states, respectively. Mechanical examples of a quadratic form as energy and work are shown in Section 3.2.1. Hence, estimation and prediction of quantities represented by an NFS can be helpful in different applications, such as system control and target tracking.

3. Theoretical Results: Optimal Mean Square Estimator for NFS

3.1. Main Idea and General Solution

The main idea of the proposed optimal estimator includes two stages: the optimal Kalman estimate of a state vector computed at the first stage is nonlinearly transformed at the second stage based on the NFS (2) and the MMSE criterion.

At first, according to the MMSE criterion, the optimal estimate of the state x_k represents the conditional mean, i.e., $\hat{x}_k = \mathbf{E}(x_k|Y^k)$. The conditional mean and corresponding covariance $P_k = \text{Cov}(e_k, e_k)$, $e_k = x_k - \hat{x}_k$, are described by the recursive Kalman filter (KF) equations [2,8–10]:

$$\begin{aligned}\hat{x}_{k+1}^- &= F_k \hat{x}_k, \quad k = 0, 1, \dots; \quad \hat{x}_0 = m_0, \\ P_{k+1}^- &= F_k P_k F_k^T + G_k Q_k G_k^T, \quad P_0 = C_0, \\ K_{k+1} &= P_{k+1}^- H_{k+1}^T (H_{k+1} P_{k+1}^- H_{k+1}^T + R_{k+1})^{-1}, \\ \hat{x}_{k+1} &= \hat{x}_{k+1}^- + K_{k+1} (y_{k+1} - H_{k+1} \hat{x}_{k+1}^-), \\ P_{k+1} &= (I_n - K_{k+1} H_{k+1}) P_{k+1}^-.\end{aligned}\quad (4)$$

Next, the optimal mean square estimate of the NFS $z_k = f(x_k)$ based on the overall sensor measurements Y^k , also represents a conditional mean:

$$\hat{z}_k = \mathbf{E}(z_k|Y^k) = \int f(x) p_k(x|Y^k) dx, \quad (5)$$

where $p_k(x|Y^k) = \mathcal{N}(\hat{x}_k, P_k)$ is a multivariate conditional Gaussian probability density function.

Thus, the best estimate in (5) represents the MMSE estimator:

$$\hat{z}_k = F(\hat{x}_k, P_k), \quad (6)$$

which depends on the Kalman estimate $\hat{x}_k$ and its error covariance P_k determined by the KF Equation (4).

Remark: For the general NFS, $z_k = f(x_k)$, calculation of the optimal estimate, $\hat{z}_k = \mathbf{E}(z_k|Y^k)$, is reduced to calculation of the multivariate integral (5). The shortcoming of the proposed estimator can be attributed to the impossibility of analytically calculating the integrals (optimal estimate) for an arbitrary class of nonlinear functionals. Analytical calculation of the integrals is possible only in special cases which are considered in Sections 3.2.2 and 3.3. In these cases, the conditional mean $\mathbf{E}(z_k|Y^k)$ can be explicitly calculated in terms of $\hat{x}_k$ and P_k , and the best MMSE estimator (5) takes a closed-form.

3.2. Optimal Closed-Form MMSE Estimator for Quadratic Functional

Consider an arbitrary quadratic functional (QF) of a state vector $x_k \in \mathbb{R}^n$:

$$z_k = f(x_k) = x_k^T \Omega_k x_k + a_k^T x_k, \quad (7)$$

where $\Omega_k = \Omega_k^T \geq 0$ and a_k are an arbitrary matrix and vector, respectively.

3.2.1. Examples of Quadratic Functional as Energy and Work

Example 1. Work done from a moving particle

Consider a particle with mass M and speed $v_k = v(t_k)$, which is moving along the x -axis. Let the particle have a velocity v_k at displacement $x_k = x(t_k)$. Then current work done, $z_k = W_k$, represents a QF such as:

$$\begin{aligned}W_k &= \frac{1}{2} M v_k^2 - \frac{1}{2} M v_0^2 = X_k^T \Omega X_k, \\ \text{where} \quad X_k &= \begin{bmatrix} v_k \\ v_0 \end{bmatrix}, \quad \Omega = \begin{bmatrix} -M & 0 \\ 0 & M \end{bmatrix}.\end{aligned}\quad (8)$$

Example 2. Energy of a harmonic oscillator

The equation of motion of a harmonic oscillator in a random environment is:

$$\begin{aligned}M \ddot{x}_t &= -k x_t + \xi_t, \\ \text{or} \quad \begin{bmatrix} \dot{x}_t \\ \dot{v}_t \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ -\omega^2 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ v_t \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \xi_t,\end{aligned}\quad (9)$$

where $v_t = \dot{x}_t$, $\omega^2 = k/M$, M is the mass, k is the force constant of a spring. White Gaussian noise ξ_t is added to represent the random environment.

Let sampling period be $\Delta t = t_k - t_{k-1} \ll 1$ so that terms of order Δt^2 can be ignored. Then the discretized model of the harmonic oscillator is approximately:

$$X_{k+1} = \begin{bmatrix} 1 & \Delta t \\ -\omega^2 \Delta t & 1 \end{bmatrix} X_k + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tilde{\xi}_k, \quad X_k = \begin{bmatrix} x_k \\ v_k \end{bmatrix}, \quad (10)$$

where $\tilde{\xi}_k$ represents discrete white noise.

Then energy, $z_k = E_k$, is conserved as:

$$E_k = \frac{1}{2} M v_k^2 + \frac{1}{2} k x_k^2 = \frac{1}{2} X_k^T \Omega X_k, \quad \Omega = \begin{bmatrix} k & 0 \\ 0 & M \end{bmatrix}. \quad (11)$$

Here the potential energy is $U_k = kx_k^2/2$.

3.2.2. Closed-Form Quadratic Estimator

In the case of the QF (7) the optimal nonlinear estimate (5) can be explicitly calculated in terms of a state estimate $\hat{x}_k$ and its error covariance P_k . Using the formula [30]:

$$\begin{aligned} \mathbf{E}(x^T \Omega x) &= \text{tr}[\Omega(C + mm^T)], \\ m &= \mathbf{E}(x), \quad C = \text{Cov}(x, x), \end{aligned} \quad (12)$$

we obtain the optimal mean-square estimate for the QF:

$$\begin{aligned} \hat{z}_k^{\text{opt}} &= \mathbf{E}(x_k^T \Omega_k x_k + a_k^T x_k | Y^k) \\ &= \text{tr}[\Omega_k(P_k + \hat{x}_k \hat{x}_k^T)] + a_k^T \hat{x}_k. \end{aligned} \quad (13)$$

In (12) and (13), $\text{tr}(A)$ is the trace of a matrix A .

In parallel to the optimal estimate (13) consider a suboptimal estimate of the QF, such as $\hat{z}_k^{\text{sub}} = f(\hat{x}_k)$:

$$\hat{z}_k^{\text{sub}} = \hat{x}_k^T \Omega_k \hat{x}_k + a_k^T \hat{x}_k. \quad (14)$$

3.2.3. True Mean Square Error for Quadratic Estimators

We compare the estimation accuracy of the optimal and suboptimal quadratic estimators (13) and (14), respectively.

The following theorem defines the true mean square errors (MSEs):

$$\begin{aligned} P_{z,k}^{\text{opt}} &= \mathbf{E}[(z_k - \hat{z}_k^{\text{opt}})^2], \\ P_{z,k}^{\text{sub}} &= \mathbf{E}[(z_k - \hat{z}_k^{\text{sub}})^2]. \end{aligned} \quad (15)$$

Theorem 1. The mean square errors $P_{z,k}^{\text{opt}}$ and $P_{z,k}^{\text{sub}}$ for the optimal and suboptimal quadratic estimators (13) and (14) are given by:

$$\begin{aligned} P_{z,k}^{\text{opt}} &= 4\text{tr}(\Omega_k P_k \Omega_k C_k) - 2\text{tr}(\Omega_k P_k \Omega_k P_k) \\ &\quad + 4\mu_k^T \Omega_k P_k \Omega_k \mu_k + a_k^T P_k a_k + 4\mu_k^T \Omega_k P_k a_k, \end{aligned} \quad (16)$$

and

$$\begin{aligned} P_{z,k}^{\text{sub}} &= 4\text{tr}(\Omega_k P_k \Omega_k C_k) - 2\text{tr}(\Omega_k P_k \Omega_k P_k) \\ &\quad + 4\mu_k^T \Omega_k P_k \Omega_k \mu_k + \text{tr}^2(\Omega_k P_k^-) \\ &\quad + a_k^T P_k a_k + 4\mu_k^T \Omega_k P_k a_k, \end{aligned} \quad (17)$$

respectively. Here the mean $\mu_k = \mathbf{E}(x_k)$ and covariance $C_k = \text{Cov}(x_k, x_k)$ are determined by the Lyapunov equation:

$$\begin{aligned} \mu_{k+1} &= F_k \mu_k, \quad k = 0, 1, \dots; \quad \mu_0 = m_0 = \mathbf{E}(x_0), \\ C_{k+1} &= F_k C_k F_k^T + G_k Q_k G_k^T, \quad C_0 = \text{Cov}(x_0, x_0). \end{aligned} \quad (18)$$

The proof of the theorem is given in the Appendix A.

Corollary 1. In the particular case with $z_k = \|x_k\|^2 = x_k^T x_k$, the optimal and suboptimal quadratic estimators and MSEs take the form:

$$\begin{aligned}\hat{z}_k^{\text{opt}} &= \text{tr}(P_k + \hat{x}_k \hat{x}_k^T) = \hat{x}_k^T \hat{x}_k + \text{tr}(P_k), \\ \hat{z}_k^{\text{sub}} &= \hat{x}_k^T \hat{x}_k, \\ P_{z,k}^{\text{opt}} &= 4\text{tr}(P_k C_k) - 2\text{tr}(P_k^2) + 4\mu_k^T P_k \mu_k, \\ P_{z,k}^{\text{sub}} &= 4\text{tr}(P_k C_k) - \text{tr}(P_k^2) + 4\mu_k^T P_k \mu_k.\end{aligned}\quad (19)$$

Generalization of a QF is a multivariate polynomial functional, which we consider in the next section.

3.3. Optimal Closed-Form MMSE Estimator for Polynomial Functional

Here we consider a special NFS which represents an arbitrary multivariate polynomial functional of an Nth degree:

$$z_k = f(x_k) = \sum_{0 \leq \ell_1 + \dots + \ell_n \leq N} A_{\ell_1 \ell_2 \dots \ell_n} x_{1,k}^{\ell_1} x_{2,k}^{\ell_2} \dots x_{n,k}^{\ell_n}, \quad (20)$$

$$\ell_1, \ell_2, \dots, \ell_n = 0, 1, \dots, N,$$

where $A_{\ell_1 \ell_2 \dots \ell_n}$ are nonzero coefficients.

Further for simplicity, we ignore the subscript k of $x_k, \hat{x}_k, x_{i,k}, \hat{x}_{i,k}$ and $P_k, P_{ij,k}$. In case of the polynomial functional (20), the MMSE estimator, $\hat{z}_k = \mathbf{E}[f(x)|Y^k]$, has a closed form because the best estimate (conditional expectation)

$$\begin{aligned}\hat{z}_k &= \mathbf{E}[f(x)|Y^k] \\ &= \sum_{0 \leq \ell_1 + \dots + \ell_n \leq N} A_{\ell_1 \ell_2 \dots \ell_n} \mathbf{E}(x_1^{\ell_1} x_2^{\ell_2} \dots x_n^{\ell_n} | Y^k),\end{aligned}\quad (21)$$

depends on high-order conditional moments, $m_{\ell_1 \ell_2 \dots \ell_n} = \mathbf{E}(x_1^{\ell_1} x_2^{\ell_2} \dots x_n^{\ell_n} | Y^k)$, of a multivariate Gaussian distribution, which can be explicitly calculated in terms of first- and second-order conditional moments, namely, the Kalman estimate and its error covariance $(\hat{x}, P)$ [31,32]. For example, the third- and fourth-order moments of the components x_i of a state vector $x^T = [x_1 \dots x_n]$ are:

$$\begin{aligned}\mathbf{E}(x_i | Y^k) &= \hat{x}_i, \\ \mathbf{E}(x_i x_j | Y^k) &= \hat{x}_i \hat{x}_j + P_{ij}, \\ \mathbf{E}(x_i x_j x_\ell | Y^k) &= \hat{x}_i \hat{x}_j \hat{x}_\ell + \hat{x}_i P_{j\ell} + \hat{x}_j P_{i\ell} + \hat{x}_\ell P_{ij}, \\ \mathbf{E}(x_i x_j x_h x_\ell | Y^k) &= \hat{x}_i \hat{x}_j \hat{x}_h \hat{x}_\ell + P_{ij} P_{h\ell} + P_{ih} P_{j\ell} \\ &\quad + P_{i\ell} P_{jh} + \hat{x}_i \hat{x}_j P_{h\ell} + \hat{x}_i \hat{x}_h P_{j\ell} + \hat{x}_i \hat{x}_\ell P_{jh} \\ &\quad + \hat{x}_j \hat{x}_h P_{i\ell} + \hat{x}_j \hat{x}_\ell P_{ih} + \hat{x}_h \hat{x}_\ell P_{ij},\end{aligned}\quad (22)$$

where

$$\hat{x}^T = [\hat{x}_1 \dots \hat{x}_n], \quad P = [P_{ij}], \quad i, j = 1, \dots, n.$$

For an arbitrary NFS, $z_k = f(x_k)$, we can use a truncated Taylor series. Then an optimal estimate of NFS can be approximately evaluated using the analytical Equation (22) for high-order moments. Also, to avoid complicated calculations of the multivariate integral in (5) it is useful to apply the unscented transformation.

3.4. Application of Unscented Transformation for Estimation of General NFS

The unscented transformation (UT) approximates the statistics of a transformed random variable, for example, the mean and covariance [33]. Following this approach, the procedure to calculate the best estimate of an NFS, $\hat{z}_k^{\text{opt}} = \mathbf{E}[f(x_k)|Y^k]$, using the UT can be summarized as follows:

$$\hat{z}_k^{\text{opt}} \approx \hat{z}_k^{\text{UT}} = \sum_{s=0}^{2n} \omega^{(s)} f(X_k^{(s)}) \quad (23)$$

where the sigma points $X_k^{(s)}$ and corresponding weights $\omega_k^{(s)}$ are determined by:

$$\begin{aligned} X_k^{(0)} &= \hat{x}_k, \quad X_k^{(s)} = \hat{x}_k + \left[\sqrt{(n+\ell)P_k} \right]_s, \\ X_k^{(s+n)} &= \hat{x}_k - \left[\sqrt{(n+\ell)P_k} \right]_s, \quad \omega^{(0)} = \frac{\ell}{n+\ell}, \\ \omega^{(s)} &= \omega^{(s+n)} = \frac{1}{2(n+\ell)}, \quad s = 1, \dots, n. \end{aligned} \quad (24)$$

Here $\left[\sqrt{P_k} \right]_s$ is the s th column of the matrix square root of P_k , and ℓ is the scaling parameter [34]. Thus, the UT estimate $\hat{z}_k^{UT}$ is represented by known functional of the Kalman estimate $\hat{x}_k$ and error covariance P_k .

4. Numerical Verification

In this section, demonstrative examples are used for testing the effectiveness of the MMSE estimators developed in this paper.

4.1. Estimation of Distance between Unknown and Moving Points

If θ is a scalar unknown which is measured in the presence of additive white noise then:

$$\begin{aligned} x_{k+1} &= x_k, \quad k = 0, 1, \dots, \quad x_0 \equiv \theta \sim \mathcal{N}(\bar{\theta}, \sigma_\theta^2), \\ y_k &= x_k + w_k, \quad w_k \sim \mathcal{N}(0, r). \end{aligned} \quad (25)$$

The KF Equation (4) becomes:

$$\begin{aligned} \hat{x}_{k+1} &= \hat{x}_k + K_{k+1}(y_{k+1} - \hat{x}_k), \quad \hat{x}_0 = \bar{\theta}, \\ P_{k+1}^- &= P_k, \quad K_{k+1} = P_k / (r + P_k), \quad P_0 = \sigma_\theta^2, \\ P_{k+1} &= (1 - K_{k+1})P_{k+1}, \quad k = 0, 1, \dots \end{aligned} \quad (26)$$

Using “step-by-step” induction, we obtain an exact formula for the error covariance $P_k = \mathbf{E}[(\theta - \hat{x}_k)^2]$:

$$P_k = \frac{r\sigma_\theta^2}{r + k\sigma_\theta^2}, \quad k = 0, 1, \dots \quad (27)$$

Further, we consider an NFS representing distance between two points. In this case an NFS becomes $z_k = |\theta - a_k|$. Using (5), the best estimate of the distance between an unknown point (location) θ and the moving point $a_k, k = 1, 2, \dots$, is:

$$\begin{aligned} \hat{z}_k &= \int_{-\infty}^{\infty} |\theta - a_k| \mathcal{N}(\hat{x}_k, P_k) d\theta = \sqrt{\frac{2P_k}{\pi}} \exp\left(-\frac{\hat{x}_k^2}{2P_k}\right) \\ &\quad + \hat{x}_k \left[1 - 2\Phi\left(-\frac{\hat{x}_k}{\sqrt{P_k}}\right) \right] + a_k \left[2\Phi\left(\frac{a_k - \hat{x}_k}{\sqrt{P_k}}\right) - 1 \right], \end{aligned} \quad (28)$$

where $\hat{x}_k$ and P_k are determined by (26) and (27), respectively, and $\Phi(u)$ is the Gaussian cumulative distribution function:

$$\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-t^2/2} dt. \quad (29)$$

4.2. Estimation of Power and Modulus of Unknown Signal

If x_k is a scalar random signal measured in additive white noise then the system model is

$$\begin{aligned} x_{k+1} &= ax_k + v_k, \quad x_0 \sim \mathcal{N}(m_0, \sigma_0^2), \\ y_k &= x_k + w_k, \quad k = 1, 2, \dots, \end{aligned} \quad (30)$$

where $v_k \sim \mathcal{N}(0, q)$, and $w_k \sim \mathcal{N}(0, r)$ are the uncorrelated white Gaussian noises, and $a \in (0, 1)$.

The KF Equation (4) gives the following:

$$\begin{aligned}
 \hat{x}_{k+1} &= a\hat{x}_k + K_{k+1}(y_{k+1} - a\hat{x}_k), \quad \hat{x}_0 = m_0, \\
 P_{k+1}^- &= a^2 P_k + q, \quad K_{k+1} = \frac{P_k^-}{(r + P_k^-)}, \quad P_0 = \sigma_0^2, \\
 P_{k+1} &= (1 - K_{k+1})P_k^-, \quad k = 0, 1, \dots
 \end{aligned} \tag{31}$$

Further, we consider two specific nonlinear functionals of the scalar signal x_k . They represent power and modulus of the signal.

4.2.1. Estimation of Power of Signal

In this case an NFS represents quadratic functional, $z_k = f(x_k) = x_k^2$. Using Equation (13) we obtain the optimal MMSE estimate of power of the signal, $\hat{z}_k = \hat{x}_k^2 + P_k$. In addition to the optimal estimate consider a simple suboptimal estimate of the power, such as $\tilde{z}_k = f(\hat{x}_k) = \hat{x}_k^2$.

We can compare the estimation accuracy of the optimal and suboptimal estimates. Using Theorem 1 with Formulas (16)–(18) we derive the precise formulas for the true MSEs of these estimates, $P_k^{\text{opt}} = \mathbf{E}[(z_k - \hat{z}_k)^2]$ and $P_k^{\text{sub}} = \mathbf{E}[(z_k - \tilde{z}_k)^2]$, respectively:

$$\begin{aligned}
 P_k^{\text{opt}} &= \mathbf{E}[(x_k^2 - \hat{x}_k^2 - P_k)^2] = 4P_k C_k - 2P_k^2 + 4\mu_k^2 P_k, \\
 P_k^{\text{sub}} &= \mathbf{E}[(x_k - \hat{x}_k)^2] = 4P_k C_k - P_k^2 + 4\mu_k^2 P_k.
 \end{aligned} \tag{32}$$

Here the mean μ_k and covariance C_k of the signal x_k are determined by the Lyapunov Equation (18):

$$\mu_{k+1} = a\mu_k, \quad C_{k+1} = a^2 C_k + q. \tag{33}$$

The analytical solution of the Lyapunov equations is:

$$\mu_k = a^k m_0, \quad C_k = a^{2k} \sigma_0^2 + q \sum_{i=0}^{k-1} a^{2i}, \quad k = 1, 2, \dots \tag{34}$$

Thus, the KF Equation (31) with Formulas (32)–(34) completely establish the true MSEs for the optimal and suboptimal estimates, $\hat{z}_k = \hat{x}_k^2 + P_k$, and $\tilde{z}_k = \hat{x}_k^2$, respectively. The difference between P_k^{opt} and P_k^{sub} is equal to $P_k^{\text{sub}} - P_k^{\text{opt}} = P_k^2$. Figures 2 and 3 show the MSEs and the relative error $\Delta_k = |(P_k^{\text{sub}} - P_k^{\text{opt}})/P_k^{\text{opt}}| 100\%$ for the values $a = 0.9$, $q = 0.05$, $m_0 = 0$, $\sigma_0^2 = 4$, and $r = 1$, respectively. From Figure 3 we observe that the relative error $\Delta_k(\%)$ varies from 3% to 6% within the time zone $k \leq 11$, and then increases. In a steady-state regime, $k > 40$, the relative error reaches the value $\Delta_\infty = 18.3\%$, and at the same time zone the absolute values of the MSEs are equal to $P_\infty^{\text{opt}} = 0.1087$ and $P_\infty^{\text{sub}} = 0.1286$. Thus, the numerical results show that the suboptimal estimate $\tilde{z}_k = \hat{x}_k^2$ may be significantly worse than the optimal one $\hat{z}_k = P_k + \hat{x}_k^2$.

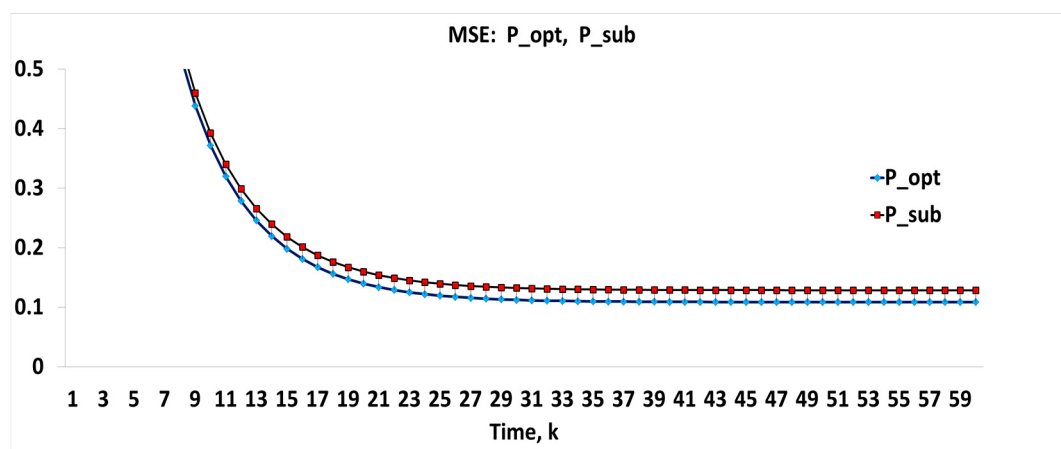


Figure 2. Optimal and suboptimal MSEs for power of signal $z_k = x_k^2$.

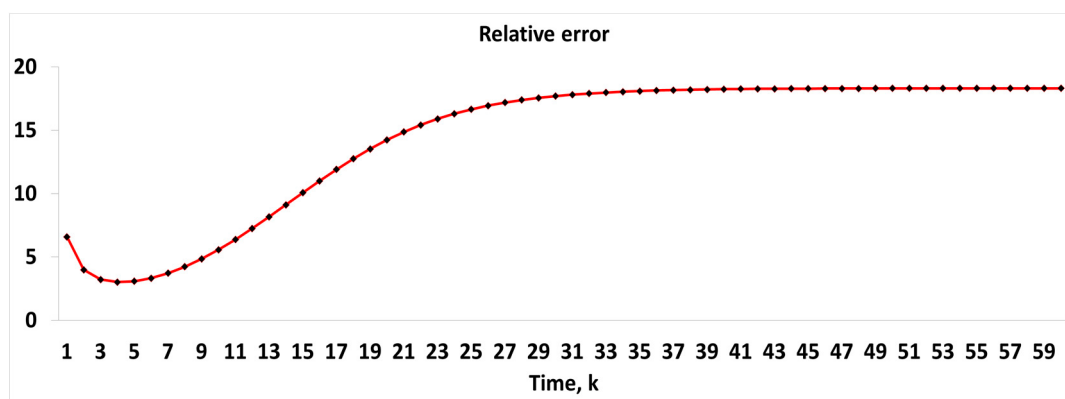


Figure 3. Relative error for power of signal $z_k = x_k^2$.

4.2.2. Estimation of Modulus of Signal

Based on Equation (28), the best estimate of an unknown modulus $z_k = f(x_k) = |x_k|$ takes the form:

$$\hat{z}_k = \sqrt{\frac{2P_k}{\pi}} \exp\left(-\frac{\hat{x}_k^2}{2P_k}\right) + \hat{x}_k \left[1 - 2\Phi\left(-\frac{\hat{x}_k}{\sqrt{P_k}}\right)\right]. \quad (35)$$

In contrast to the optimal estimate (35) we can consider a simple suboptimal estimate:

$$\tilde{z}_k = f(\hat{x}_k) = |\hat{x}_k|. \quad (36)$$

To study the behavior of the true MSEs, $P_k^{\text{opt}} = \mathbf{E}[(z_k - \hat{z}_k)^2]$ and $P_k^{\text{sub}} = \mathbf{E}[(z_k - \tilde{z}_k)^2]$, set $a = 0.99$, $q = 1$, $r = 0.5$, and $x_0 \sim \mathcal{N}(1; 4)$. To compare the MSEs, a Monte-Carlo simulation with 1000 runs was used. As shown in Figure 4, the optimal estimate $\hat{z}_k$ demonstrates an improvement over the suboptimal estimate $\tilde{z}_k$.

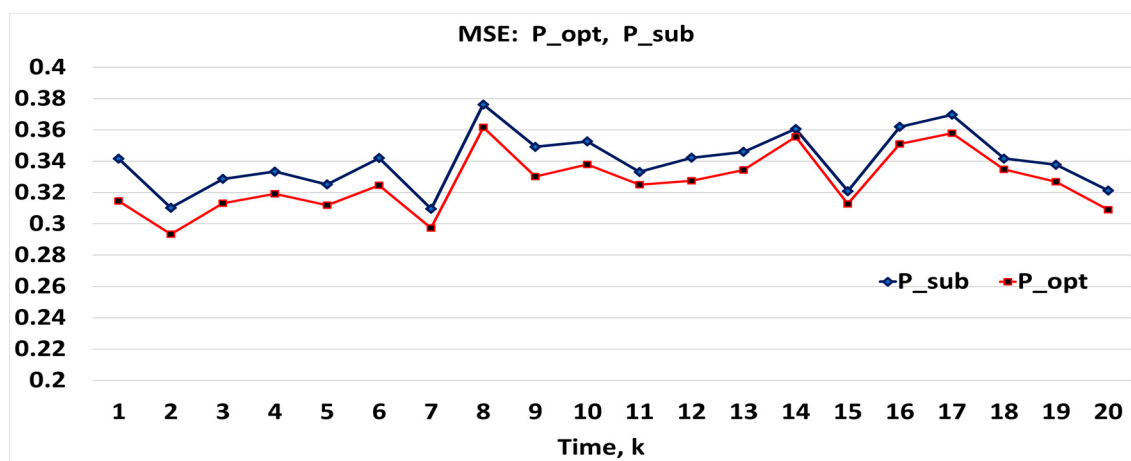


Figure 4. MSEs of estimates for modulus function $z_k = |x_k|$.

4.3. Numerical Example of QF: Wind Tunnel System

Following is a comparative experimental analysis of the optimal and suboptimal quadratic estimators considered in an example of total kinetic energy of the wind tunnel system.

A discrete-time high-speed closed-air unit wind tunnel model is given in [35]:

$$x_{k+1} = \begin{bmatrix} 0.9032 & 0 & 0 \\ 0 & 0.6245 & 0.0701 \\ 0 & -2.5247 & -0.0488 \end{bmatrix} x_k + v_k. \quad (37)$$

The state vector $x_k \in \mathbb{R}^3$ consists of the state variables $x_{1,k}$, $x_{2,k}$ and $x_{3,k}$, representing derivatives from a chosen equilibrium point of the following quantities: x_1 = Mach number, x_2 = actuator position guide vane angle in a driving fan, and x_3 = actuator rate. The initial mean and

covariance are $m_0 = [3 \ 28 \ 10]^T$ and $C_0 = \text{diag}[1 \ 1 \ 0]$, the sampling period of $\Delta t = 0.01$, and $v_k \in \mathbb{R}^3$ is white Gaussian noise, $v_k \sim \mathcal{N}(0, Q)$, $Q = \text{diag}[0.01^2 \ 0.01^2 \ 0.01^2]$.

Two sensory measurement model is given by:

$$y_k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} x_k + w_k, \quad w_k \sim \mathcal{N}(0, R), \quad (38)$$

$$R = \text{diag}[0.5 \ 0.5].$$

The total kinetic energy of an actuator represented by the QF, $z_k = x_k^T \Omega x_k$, can be expressed as the sum of the translational kinetic energy of the center of mass, $E_k^t = M v_k^2 / 2$, and the rotational kinetic energy about the center of mass, $E_k^r = I \omega_k^2 / 2$, where I is rotational inertia, $\omega_k = \dot{x}_{2,k}$ is angular velocity, M is mass and $v_k = x_{3,k}$ is linear velocity.

Applying finite difference approximation for the velocity, $\omega_k \approx (x_{2,k} - x_{2,k-1}) / \Delta t$, the total kinetic energy can be expressed in the following QF (see Section 3.2.2):

$$z_k = E_k^r + E_k^t = \frac{1}{2} I \dot{x}_{2,k}^2 + \frac{1}{2} M x_{3,k}^2 = \frac{1}{2} X_k^T \Omega X_k, \quad (39)$$

$$X_k = \begin{bmatrix} x_{1,k} \\ x_{2,k} \\ x_{3,k} \\ x_{2,k-1} \end{bmatrix}, \quad \Omega = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\Delta t^2} & 0 & -\frac{1}{\Delta t} \\ 0 & 0 & M & 0 \\ 0 & -\frac{1}{\Delta t} & 0 & \frac{1}{\Delta t^2} \end{bmatrix},$$

where $X_k \in \mathbb{R}^4$ is extended state vector, and $I = 0.136 \text{ kgm}^2$, $M = 7.39 \text{ kg}$.

Using the obtained results of Equations (13) and (14) the optimal and suboptimal quadratic estimators take the form:

$$\hat{z}_k^{\text{opt}} = \text{tr}[\Omega(P_k + \hat{X}_k \hat{X}_k^T)], \quad \hat{z}_k^{\text{sub}} = \hat{X}_k^T \Omega \hat{X}_k, \quad (40)$$

respectively. The estimate of the state $\hat{X}_k \in \mathbb{R}^4$ and error covariance $P_k \in \mathbb{R}^{4 \times 4}$ are determined by the KF Equation (4).

Our point of interest is the behavior of MSEs of the estimators (40). Using Theorem 1, the MSEs are given by

$$P_{z,k}^{\text{opt}} = 4\text{tr}(\Omega P_k \Omega C_k) - 2\text{tr}(\Omega P_k \Omega P_k) + 4\mu_k^T \Omega P_k \Omega \mu_k, \quad (41)$$

$$P_{z,k}^{\text{sub}} = P_{z,k}^{\text{opt}} + \text{tr}^2(\Omega P_k),$$

respectively. Here the mean μ_k and covariance C_k of the state vector X_k satisfy the Lyapunov Equation (18).

We observe in Figure 5 that the difference between absolute values of the MSEs P_k^{opt} and P_k^{sub} is very small; for example, $P_k^{\text{opt}} \approx 0.938$ and $P_k^{\text{sub}} \approx 0.884$ at $k > 10$. However, the relative error exceeds 6% at the same time zone, $\Delta_k(\%) \approx 6.2\%$. For this reason we conclude that the suboptimal estimator is not suitable for evaluation of the kinetic energy of the wind tunnel system.

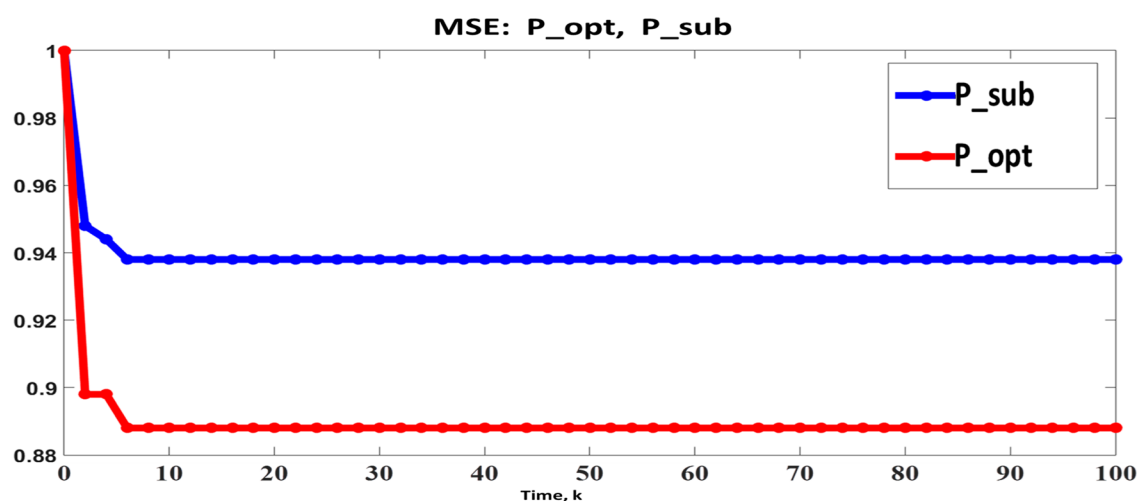


Figure 5. Comparison of MSEs of total kinetic energy z_k using optimal and suboptimal estimators.

4.4. Numerical Example of General NFS: Motion of Unmanned Marine Probe

A comparative experimental analysis of the proposed estimators is considered for the motion of an unmanned marine probe (UMP). In a marine inspection environment, an UMP system is often considered because of their benefits of convenience and human safety.

Assume a scenario in which the UMP detected an oil-tanker accident, from which oil has spread on the surface of the water without the influence of wind. As an initial action, the UMP estimates the length of a contour of the oil spread (Figure 6).

To control the size of the surface, the UMP must compute the distance from the oil-tanker $d_k = d_{t_k}$ at every time instance represented by the NFS:

$$d_k = f(x_k) = (x_{1,k}^2 + x_{2,k}^2)^{1/2}, \quad (42)$$

where $x_{1,k}$ and $x_{2,k}$ are coordinates of UMP.

Here, we verify the proposed estimators using a discretization of a continuous-time model of the UMP [36]:

$$\begin{aligned} \dot{x}_{1,t} &= x_{1,t} - 2x_{2,t} + v_{1,t}, \quad t \in [0; 3], \\ \dot{x}_{2,t} &= x_{1,t} - x_{2,t} + v_{2,t}, \end{aligned} \quad (43)$$

where $v_{1,t}$ and $v_{2,t}$ are uncorrelated white Gaussian noises with intensities $q_1 = q_2 = 0.1$; $x_{1,0} \sim \mathcal{N}(20; 0.2)$, and $x_{2,0} \sim \mathcal{N}(0; 0.2)$.

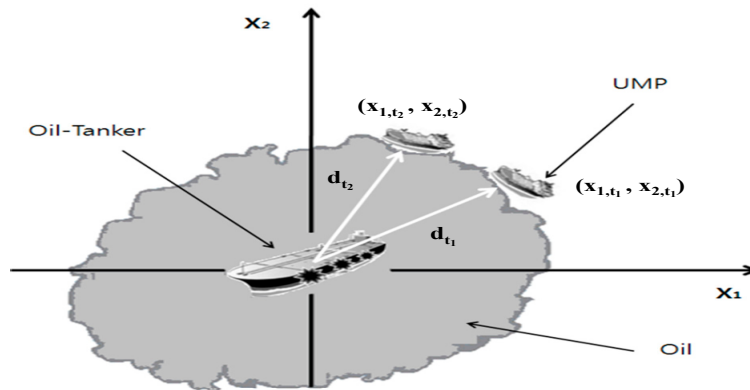


Figure 6. Estimation of size of oil spread contour.

The UMP measures the relative coordinates $x_{1,k}$ and $x_{2,k}$ from the oil-tanker, respectively. The measurement model for the UMP is given by:

$$\begin{aligned} y_{1,k} &= x_{1,k} + w_{1,k}, \\ y_{2,k} &= x_{2,k} + w_{2,k}, \end{aligned} \quad (44)$$

where $w_{1,k} \sim \mathcal{N}(0; 0.1)$ and $w_{2,k} \sim \mathcal{N}(0; 0.1)$ are uncorrelated white Gaussian noises.

A comparative experimental analysis of the optimal estimate $\hat{d}_k^{\text{opt}}$ which is based on the UT (23) and the suboptimal estimate, $\hat{d}_k^{\text{sub}} = f(\hat{x}_k) = (\hat{x}_{1,k}^2 + \hat{x}_{2,k}^2)^{1/2}$, is considered. To compare the MSEs $p_k^{\text{opt}} = \mathbf{E}[(d_k - \hat{d}_k^{\text{opt}})^2]$ and $p_k^{\text{sub}} = \mathbf{E}[(d_k - \hat{d}_k^{\text{sub}})^2]$, a Monte-Carlo simulation with 1000 runs and finite-difference approximation with the step $\Delta t = 0.01$ were performed. Figure 7 illustrates the time histories of the MSEs for the both estimators. It shows that the optimal estimate $\hat{d}_k^{\text{opt}}$ has the best performance due to the lowest value of the MSE, $p_k^{\text{opt}} < p_k^{\text{sub}}$.

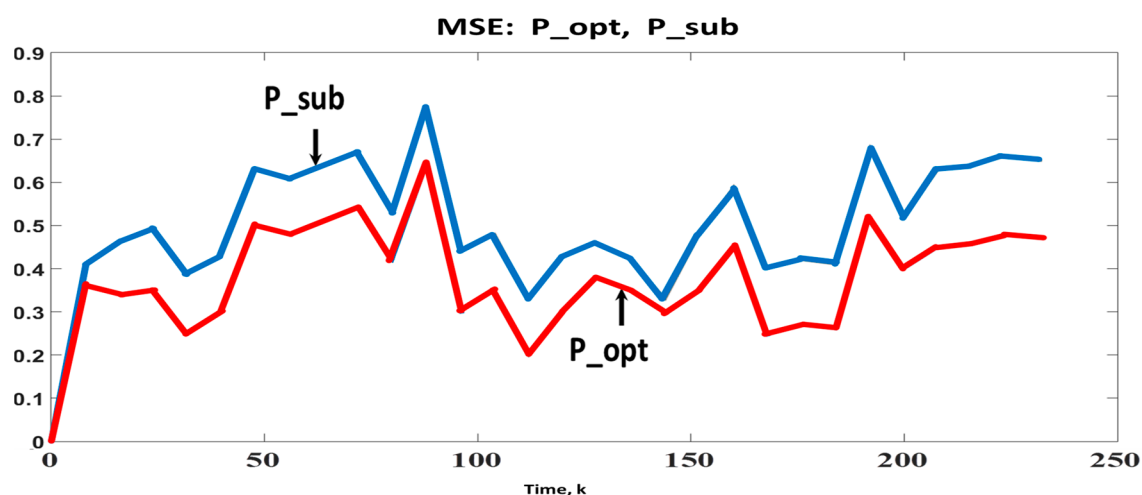


Figure 7. MSEs of optimal and suboptimal estimators for oil spread contour d_k .

5. Conclusions

In some application problems, nonlinear functional of state variables are interpreted as a cost function that denotes useful information of a target system for control. In order to estimate an arbitrary NFS, the optimal MMSE estimator is proposed. However, calculation of the optimal estimate $\hat{z}_k^{\text{opt}}$ is reduced to calculation of the multivariate integral (Equation (5)) or usage of the unscented transformation $\hat{z}_k^{\text{UT}}$. To avoid these numerical difficulties, two important classes of an NFS (quadratic and polynomial) are studied in detail. Analytical calculation of the integral (Equation (5)) in terms of the Kalman statistics $(\hat{x}_k, P_k)$ is possible for these functionals and, as a result, effective closed-form quadratic and polynomial MMSE estimators are derived.

Special attention is given for quadratic functionals. In this case, the quadratic estimator is comprehensively investigated, and novel compact matrix forms for the optimal estimate (Equation (13)) and true MSE (Theorem 1) are derived.

In view of the importance of an NFS in practice, the proposed estimation algorithms are illustrated on theoretical and practical examples for a real NFS. The examples show that the optimal estimator yields reasonably good estimation accuracy, and we confirm that the proposed optimal MMSE estimator is suitable for data processing in practice although, in some situations, the suboptimal estimators are slightly worse (in terms of absolute MSE) than the optimal estimator (Sections 4.2.1 and 4.3). In such situations, we propose an additional comparison of the estimators through a relative error $\Delta_k(\%)$, which exceeds 6% and demonstrates the advantage of the optimal estimator over the suboptimal one.

The KF gain K_k and error covariance P_k may be pre-computed, since they do not depend on current measurements but only on the noise statistics and system matrices, which are part of the system model (Equation (1)). Thus, once the measurement schedule has been settled, real-time implementation of the MMSE estimator (Equation (6)) or (Equation (13)) requires only the computation of the state estimate and final estimate of the NFS.

The optimality and statistical accuracy of the proposed estimator of NFS as well as the KF depends on several factors, such as quality of prior assumptions about the system model (Equation (1)), especially the process Q_k and measurement R_k noise covariances, respectively. The quality of prior assumptions is an important factor that leads to the optimality of estimates. Inadequacy of prior information could lead to unexpected results and estimator divergence. Determination of suitable values for noise covariances and system matrices plays a crucial role in obtaining a converged estimator. Adaptive filtering is one of the powerful approaches to prevent the divergence problem of the filter when precise knowledge on the prior assumptions is not available [37,38]. Therefore, the adaptive Kalman estimate $\hat{x}_k$ can be used in the proposed MMSE estimator (6).

For future work, we propose to extend the MMSE estimator for nonlinear functionals which depend not only on the current state x_k , but also on its past $x_k, x_{k-1}, \dots, x_1$, for example, $f = \sum_{t=1}^k x_t^T Q_t x_t$.

Author Contributions: W.C. developed the theoretical formalism, performed the analytic calculations, I.Y.S. and V.S. designed and performed the numerical simulations, derived the models and analysed the data. All authors equally contributed in the elaboration of this work.

Funding: This research was supported by the Incheon National University Research Grant in 2017–2018, grant number 2017-0022.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Appendix A

The Proof of Theorem 1. The derivation of the optimal MSE (16) is based on the following lemma.

Lemma A1: Let $X \in \mathbb{R}^{3n}$ be a composite multivariate Gaussian vector, $X^T = [U^T \ V^T \ W^T]$, i.e.,

$$X \sim \mathcal{N}(x; \mu, S), \quad U, V, W \in \mathbb{R}^n, \quad (A1)$$

$$\mu = \begin{bmatrix} \mu_u \\ \mu_v \\ \mu_w \end{bmatrix}, \quad S = \begin{bmatrix} S_{uu} & S_{uv} & S_{uw} \\ S_{vu} & S_{vv} & S_{vw} \\ S_{wu} & S_{wv} & S_{ww} \end{bmatrix}.$$

Then the third- and fourth-order vector moments of the composite random vector X are given by:

$$\begin{aligned} \text{(i)} \quad \mathbf{E}(U^T V W) &= \mu_u^T \mu_v \mu_w + \text{tr}(S_{uv}) \mu_w^T + \mu_v^T S_{uw} + \mu_u^T S_{vw}; \\ \text{(ii)} \quad \mathbf{E}(U^T U V^T V) &= \mu_u^T \mu_u \mu_v^T \mu_v + 2\text{tr}(S_{uv} S_{vu}) + \text{tr}(S_{uu}) \text{tr}(S_{vv}) + \text{tr}(S_{uu}) \mu_v^T \mu_v + \text{tr}(S_{vv}) \mu_u^T \mu_u + 4\mu_u^T S_{uv} \mu_v; \\ \text{(iii)} \quad \mathbf{E}(U^T V V^T U) &= \mu_u^T \mu_v \mu_v^T \mu_u + \text{tr}(S_{uu} S_{vv}) + \text{tr}(S_{uv}) \text{tr}(S_{vu}) + \text{tr}(S_{uv}^2) + \mu_v^T S_{uu} \mu_v + \mu_u^T S_{vv} \mu_u + \mu_v^T S_{uv} \mu_u + \mu_u^T S_{vu} \mu_v + 2\text{tr}(S_{uv}) \mu_u^T \mu_v; \\ \text{(iv)} \quad \mathbf{E}(U^T V W^T U) &= \mu_u^T \mu_v \mu_w^T \mu_u + \text{tr}(S_{uv}) \text{tr}(S_{uw}) + \text{tr}(S_{uu} S_{wv}) + \text{tr}(S_{uw} S_{uv}) + \text{tr}(S_{uv}) \mu_u^T \mu_w + \text{tr}(S_{uw}) \mu_u^T \mu_v + \mu_v^T S_{uu} \mu_w + \mu_v^T S_{uw} \mu_u + \mu_u^T S_{vu} \mu_w + \mu_u^T S_{vw} \mu_u. \end{aligned} \quad (A2)$$

The derivation of the vector Formula (A2) is based on their scalar versions [31,32]:

$$\begin{aligned} \mathbf{E}(x_i x_j x_k) &= \mu_i \mu_j \mu_k + \mu_i S_{jk} + \mu_j S_{ik} + \mu_k S_{ij}; \\ \mathbf{E}(x_i x_j x_k x_\ell) &= \mu_i \mu_j \mu_k \mu_\ell + S_{ij} S_{k\ell} + S_{ik} S_{\ell j} + S_{i\ell} S_{jk} + \mu_i \mu_j S_{k\ell} + \mu_i \mu_k S_{j\ell} + \mu_i \mu_\ell S_{jk} + \mu_j \mu_k S_{i\ell} + \mu_j \mu_\ell S_{ik} + \mu_k \mu_\ell S_{ij}, \\ \text{where} \quad \mu_h &= \mathbf{E}(x_h), \quad S_{pq} = \text{Cov}(x_p, x_q), \end{aligned} \quad (A3)$$

and standard matrix manipulations.

This completes the proof of Lemma A1. $\square$

To derive Equation (16), for simplicity, we omit the time index k . Then using Equation (7) and Equation (13), the error can be written as:

$$\begin{aligned}
e_z &= z - \hat{z}^{\text{opt}} = x^T \Omega x + a^T x - \text{tr}[\Omega(P + \hat{x}\hat{x}^T)] - \\
&a^T x = x^T \Omega x - \hat{x}^T \Omega \hat{x} - \text{tr}(\Omega P) + a^T e = \\
&(e + \hat{x})^T \Omega (e + \hat{x}) - \hat{x}^T \Omega \hat{x} - \text{tr}(\Omega P) + a^T e = \\
&e^T \Omega e + 2e^T \Omega \hat{x} + a^T e - \text{tr}(\Omega P), \\
&\text{where} \\
&e = x - \hat{x}, \quad \text{tr}(\Omega \hat{x}\hat{x}^T) = \hat{x}^T \Omega \hat{x}, \quad \hat{x}^T \Omega e = e^T \Omega \hat{x}.
\end{aligned} \tag{A4}$$

Next, using the unbiased and orthogonality properties of the Kalman estimate $\mathbf{E}(e) = \mathbf{E}(e\hat{x}^T) = 0$, we obtain:

$$\begin{aligned}
P_z^{\text{opt}} &= \mathbf{E}(e_z^2) = \mathbf{E}(e^T \Omega e e^T \Omega e) + 4\mathbf{E}(e^T \Omega \hat{x} \hat{x}^T \Omega e) \\
&+ a^T P a + \text{tr}^2(\Omega P) + 4\mathbf{E}(e^T \Omega e e^T \Omega \hat{x}) + \\
&2\mathbf{E}(e^T \Omega e e^T) a - 2\text{tr}^2(\Omega P) + 4\mathbf{E}(e^T \Omega \hat{x} e^T) a.
\end{aligned} \tag{A5}$$

Using Lemma A1 we can calculate high-order moments in Equation (A5). We have:

$$\begin{aligned}
(a) \quad &\mathbf{E}(e^T \Omega e e^T \Omega e) = 2\text{tr}(\Omega P \Omega P) + \text{tr}^2(\Omega P), \\
&U = e, V = \Omega e; \\
(b) \quad &\mathbf{E}(e^T \Omega \hat{x} \hat{x}^T \Omega e) = \text{tr}(P \Omega P_{\hat{x}\hat{x}} \Omega) + \mu^T \Omega P \Omega \mu = \\
&\text{tr}(P \Omega C \Omega) - \text{tr}(\Omega P \Omega P) + \mu^T \Omega P \Omega \mu, \\
&U = e, V = \Omega \hat{x}, \\
(c) \quad &\mathbf{E}(e^T \Omega e e^T \Omega \hat{x}) = \mathbf{E}(e^T \Omega e \hat{x}^T \Omega e) = 0, \\
&U = e, V = \Omega e, W = \Omega \hat{x}, \\
(d) \quad &\mathbf{E}(e^T \Omega e e^T) = 0, \quad U = e, \quad V = \Omega e, \quad W = e, \\
(e) \quad &\mathbf{E}(e^T \Omega \hat{x} e^T) = \mu^T \Omega P, \quad U = e, V = \Omega \hat{x}, W = e, \\
&\text{where} \\
&\mu = \mathbf{E}(x) = \mathbf{E}(\hat{x}), \quad \mathbf{E}(\Omega \hat{x}) = \Omega \mu, \\
&P = \text{Cov}(e, e), \quad \text{Cov}(\Omega e, \Omega e) = \Omega P \Omega, \\
&C = \text{Cov}(x, x), \quad P_{\hat{x}\hat{x}} = \text{Cov}(\hat{x}, \hat{x}) = C - P.
\end{aligned} \tag{A6}$$

Substituting Equation (A6) into Equation (A5), and after some manipulations, we obtain the optimal MSE (16).

In the case of the suboptimal estimate $\hat{z}_k^{\text{sub}}$, the derivation of the MSE (Equation (17)) is similar.

References

1. Pappalardo, C.M.; Guida, D. Use of the Adjoint Method for Controlling the Mechanical Vibrations of Nonlinear Systems. *Machines* **2018**, *6*, 19.
2. Bryson, A.E. *Applied Optimal Control: Optimization, Estimation and Control*; Routledge: London, UK, 2018.
3. Davari, N.; Gholami, A. An Asynchronous Adaptive Direct Kalman Filter Algorithm to Improve Underwater Navigation System Performance. *IEEE Sens. J.* **2017**, *17*, 1061–1068.
4. Kulikova, M.V.; Tsyganova, J.V. Improved Discrete-Time Kalman Filtering within Singular Value Decomposition. *IET Control Theory Appl.* **2017**, *11*, 2412–2418.
5. Shmaliy, Y.; Zhao, S.; Ahn, C.K. Unbiased Finite Impulse Response Filtering: An Iterative Alternative to Kalman Filtering Ignoring Noise and Initial Conditions. *IEEE Control Syst.* **2017**, *37*, 70–89.
6. Grewal, M.S.; Andrews, A.P.; Bartone, C.G. *Global Navigation Satellite Systems, Inertial Navigation, and Integration*, 3rd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2013.
7. Rigatos, G.G.; Siano, P. Sensorless Control of Electric Motors with Kalman Filters: Applications to Robotic and Industrial Systems. *Int. J. Adv. Robot. Syst.* **2011**, *8*, 62–80.
8. Simon, D. *Optimal State Estimation. Kalman, H-Infinity, and Nonlinear Approaches*; John Wiley & Sons: Hoboken, NJ, USA, 2006.
9. Haykin, S. *Kalman Filtering and Neural Networks*; John Wiley & Sons: Hoboken, NJ, USA, 2004; Volume 47.
10. Bar-Shalom, Y.; Li, Y.; Kirubarajan, T. *Estimation with Applications to Tracking and Navigation*; John Wiley & Sons: New York, NY, USA, 2001.
11. Cai, T.T.; Low, M.G. Optimal Adaptive Estimation of a Quadratic Functional. *Ann. Stat.* **2006**, *34*, 2298–2325.

12. Robins, J.; Li, L.; Tchetgen, E.; Vaart, A. Higher Order Influence Functions and Minimax Estimation of Nonlinear Functionals. *Probab. Stat.* **2008**, *2*, 335–421.
13. Jiao, J.; Venkat, K.; Han, Y.; Weissman, T. Minimax Estimation of Functionals of Discrete Distributions. *IEEE Trans. Inf. Theory* **2015**, *61*, 2835–2885.
14. Jiao, J.; Venkat, K.; Han, Y.; Weissman, T. Maximum Likelihood Estimation of Functionals of Discrete Distributions. *IEEE Trans. Inf. Theory* **2017**, *63*, 6774–6798.
15. Amemiya, Y.; Fuller, W.A. Estimation for the Nonlinear Functional Relationship. *Ann. Stat.* **1988**, *16*, 147–160.
16. Donoho, D.L.; Nussbaum, M. Minimax Quadratic Estimation of a Quadratic Functional. *J. Complex.* **1990**, *6*, 290–323.
17. Grebenkov, D.S. Optimal and Suboptimal Quadratic Forms for Noncentered Gaussian Processes. *Phys. Rev.* **2013**, *E88*, 032140.
18. Laurent, B.; Massart, P. Adaptive Estimation of a Quadratic Functional by Model Selection. *Ann. Stat.* **2000**, *28*, 1302–1338.
19. Vladimirov, I.G.; Petersen, I.R. Directly Coupled Observers for Quantum Harmonic Oscillators with Discounted Mean Square Cost Functionals and Penalized Back-Action. In Proceedings of the IEEE Conference on Norbert Wiener in the 21st Century, Melbourne, Australia, 13–15 July 2016; pp. 78–83.
20. Sricharan, K.; Raich, R.; Hero, A.O. Estimation of Nonlinear Functionals of Densities with Confidence. *IEEE Trans. Inf. Theory* **2012**, *58*, 4135–4159.
21. Wisler, A.; Berisha, V.; Spanias, A.; Hero, A.O. Direct Estimation of Density Functionals Using a Polynomial Basis. *IEEE Trans. Signal Process.* **2018**, *66*, 558–588.
22. Taniguchi, M. On Estimation of Parameters of Gaussian Stationary Processes. *J. Appl. Probab.* **1979**, *16*, 575–591.
23. Zhao-Guo, C.; Hanman, E.J. The Distribution of Periodogram Ordinates. *J. Time Ser. Anal.* **1980**, *1*, 73–82.
24. Janas, D.; Sachs, R. Consistency for Nonlinear Functions of the Periodogram of Tapered Data. *J. Time Ser. Anal.* **1995**, *16*, 585–606.
25. Fay, G.; Moulines, E.; Soulier, P. Nonlinear Functionals of the Periodogram. *J. Time Ser. Anal.* **2002**, *23*, 523–553.
26. Noviello, C.; Fornaro, G.; Braca, P.; Martorella, M. Fast and Accurate ISAR Focusing Based on a Doppler Parameter Estimation Algorithm. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 349–353.
27. Wu, Y.; Yang, P. Minimax Rates of Entropy Estimation on Large Alphabets via Best Polynomial Approximation. *IEEE Trans. Inf. Theory* **2016**, *62*, 3702–3720.
28. Wu, Y.; Yang, P. Optimal Entropy Estimation on Large Alphabets via Best Polynomial Approximation. In Proceedings of the IEEE International Symposium on Information Theory, Hong Kong, China, 14–19 June 2015; pp. 824–828.
29. Song, I.Y.; Shevlyakov, G.; Shin, V. Estimation of Nonlinear Functions of State Vector for Linear Systems with Time-Delays and Uncertainties. *Math. Probl. Eng.* **2015**, *2015*, 217253.
30. Pugachev, V.S.; Sinityn, I.N. *Stochastic Differential Systems. Analysis and Filtering*; Wiley & Sons: New York, NY, USA, 1987.
31. Kan, R. From Moments of Sum to Moments of Product. *J. Multivar. Anal.* **2008**, *99*, 542–554.
32. Holmquist, B. Expectations of Products of Quadratic Forms in Normal Variables. *Stoch. Anal. Appl.* **1996**, *14*, 149–164.
33. Julier, S.J.; Uhlmann, J.K. Unscented Filtering and Nonlinear Estimation. *Proc. IEEE* **2004**, *92*, 401–422.
34. Chang, L.; Hu, B.; Li, A.; Qin, F. Transformed Unscented Kalman Filter. *IEEE Trans. Autom. Control* **2013**, *58*, 252–257.
35. Armstrong, E.S.; Tripp, J.S. *An Application of Multivariable Design Techniques to the Control of the National Transonic Facility*; NASA Tech. Paper 1887; NASA Langley Research Center, Hampton, VA, USA, 1981.
36. Mutambara, A.G.O. *Decentralized Estimation and Control for Multisensor Systems*; CRC Press: Boca Raton, FL, USA, 1998.
37. Chen, G. *Approximate Kalman Filtering*; World Scientific Publishing: Singapore, 1993.
38. Jiang, C.; Zhang, S.B.; Zhang, Q.Z. Adaptive Estimation of Multiple Fading Factors for GPS/INS Integrated Navigation Systems. *Sensors* **2017**, *17*, 1254.

