Automatica 82 (2017) 158–164

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Feedback quadratic filtering ☆

Filippo Cacace [a], Francesco Conte [b], Alfredo Germani [c], Giovanni Palombo [d]

[a] Università Campus Bio-Medico di Roma, Via Álvaro del Portillo, 21, 00128 Roma, Italy

[b] DITEN, Università degli studi di Genova, Via all'Opera Pia 11A, 16145, Genova, Italy

[c] DISIM, Università dell'Aquila, Via Vetoio, 67100 L'Aquila, Italy

[d] IASI CNR, Roma, Italy

ARTICLE INFO

Article history:
Received 17 March 2016
Received in revised form
11 January 2017
Accepted 4 April 2017

Keywords:
Filtering theory
Non-Gaussian processes
Kalman filters

ABSTRACT

This paper concerns the state estimation problem for linear discrete-time non-Gaussian systems. It is known that filters based on quadratic functions of the measurements processes (Quadratic Filter) improve the estimation accuracy of the optimal linear filter. In order to enlarge the class of systems, which can be processed by a Quadratic Filter, we rewrite the system model by introducing an output injection term. The resulting filter, named the Feedback Quadratic Filter, can be applied also to non asymptotically stable systems. We prove that the performance of the Feedback Quadratic Filter depends on the gain parameter of the output term, which can be chosen so that the estimation error is always less than or equal to the Quadratic Filter.

1. Introduction

In this paper, we study the state estimation problem for linear discrete-time non-Gaussian systems. In many applications, the widely used Gaussian assumption must be removed (see Spall, 1985, 2003 and Wu & Chen, 1993). In these cases, the conditional expectation, which gives the optimal minimum variance estimation, is the solution of an infinite dimensional problem (Zakai, 1969). Methods to approximate the state conditional probability density function include Monte Carlo methods (Arulampalam, Maskell, Gordon, & Clapp, 2002), sums of Gaussian densities (Arasaratnam, Haykin, & Elliott, 2007) and weighted sigma points (Julier & Uhlmann, 2004) among others. These general solutions can cope with nonlinearities and/or with the presence of noise outliers (Stojanovic & Nedic, 2015) or unknown parameters (Stojanovic & Nedic, 2016), and they generally have high computational cost. In the context of linear non-Gaussian systems, many research works aim at filtering algorithms that are easily computable (see Afshar, Yang, & Wang, 2012; Bilik & Tabrikian, 2010; Carravetta, Germani, & Raimondi, 1996; Gordon, Salmond, & Smith, 1993; Kassam & Thomas, 1976; Maryak, Spall, & Heydon, 2004; Picinbono & Devaut, 1988; Spall, 1995; Zhang, Kuai, Ren, Luo, & Lin, 2016 and the references therein). In the minimum variance framework a natural development is to use quadratic or polynomial functions of the observations to improve the estimation accuracy while preserving easy computability and recursion (Carravetta et al., 1996; De Santis, Germani, & Raimondi, 1995; Verriest, 1985). The suboptimal polynomial estimate is obtained by applying the KF to a system augmented with the powers of state and observations. A drawback of this approach is that the resulting augmented system is bilinear and the noise variance depends on the state variance of the original system. Thus, if the variance of the state grows unboundedly, so does the equivalent noise. As a consequence the stability of the resulting Quadratic Filter (QF) is guaranteed only for asymptotically stable systems. In this paper, we propose to use an output injection term to overcome this problem and obtain an internally stable QF. Furthermore, we show that any recursively implementable filter based on the use of powers of measurements has an error that depends on the choice of the gain of this output injection term, in contrast with the linear case. Thus, the gain can be chosen to achieve a smaller estimation error than the QF. A preliminary version of this work has been published in Cacace, Conte, Germani, and Palombo (2014), where the theoretical analysis was missing.

☆ The material in this paper was presented at the 53rd IEEE Conference on Decision and Control, December 15–17, 2014, Los Angeles, CA, USA. This paper was recommended for publication in revised form by Associate Editor Wei Xing Zheng under the direction of Editor Torsten Söderström.

E-mail addresses: f.cacace@unicampus.it (F. Cacace), fr.conte@unige.it (F. Conte), alfredo.germani@univaq.it (A. Germani), giovanni.palombo@univaq.it (G. Palombo).

http://dx.doi.org/10.1016/j.automatica.2017.04.046

2. Discussion on quadratic filtering

Consider the problem of state estimation for a discrete-time linear system with non-Gaussian noise in the form

$$x(k+1) = Ax(k) + Bu(k) + f_k, \qquad x(0) = x_0 \tag{1}$$
$$y(k) = Cx(k) + g_k \tag{2}$$

where $x(k) \in \mathbb{R}^n$, $u(k) \in \mathbb{R}^p$, $y(k) \in \mathbb{R}^q$, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times p}$ and $C \in \mathbb{R}^{q\times n}$. $\{f_k\}$ and $\{g_k\}$ are sequences of non-Gaussian random variables with values in $\mathbb{R}^n$ and $\mathbb{R}^q$, respectively. The system is assumed fully observable, i.e. rank $\mathcal{O}(A, C) = n$, where $\mathcal{O}(A, C)$ is the observability matrix of the pair (A, C).

Throughout the paper we use the following notations. $x^{[i]}$ is the ith Kronecker power of a vector x. $st_q^{-1}(\cdot)$ is the inverse of the stack function, which transforms a vector in $\mathbb{R}^{q\cdot l}$ into a $q \times l$ matrix (see Carravetta et al., 1996; De Santis et al., 1995). Given a random vector $x \in \mathbb{R}^i$, $\psi_x^{(i)} = E[(x - E[x])^{[i]}]$, that is, its centered ith moment.

The random sequences $\{f_k\}$ and $\{g_k\}$ and x_0 satisfy the following conditions for $k \geq 0$:

(1) $x_0 \sim \mathcal{N}(\bar{x}_0, \Psi_{x_0})$;
(2) $\{f_k\}$ and $\{g_k\}$ are sequences of zero mean temporally independent random vectors;
(3) $\{f_k\}$, $\{g_k\}$ and x_0 are statistically independent;
(4) x_0, f_k and g_k have finite fourth moments;
(5) $\psi_{x_0}^{(i)}$, $\psi_f^{(i)}$ and $\psi_g^{(i)}$, $i = 2, 3, 4$, are known vectors;
(6) $[C\ \Psi_g]$, $\Psi_g = st_q^{-1}(\psi_g^{(2)})$, is full row rank (FRR).

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\mathcal{G}$ be a given sub σ-algebra of $\mathcal{F}$ and $L^2(\mathcal{G}, n)$ be the Hilbert space of the n-dimensional, $\mathcal{G}$-measurable random variables with finite second moment. We write $L^2(X, n)$ to denote $L^2(\sigma(X), n)$, where $\sigma(X)$ is the σ-algebra generated by X. $\Pi[\cdot|\mathcal{M}]$ is the orthogonal projection onto a given Hilbert space $\mathcal{M}$. Given system (1)–(2), the output sequence $Y_k = \mathrm{col}(y(0), \ldots, y(k))$ and the auxiliary vector $Y'_k = \mathrm{col}(1, Y_k) \in \mathbb{R}^{l+1}$, $l = (k+1)q$, the minimum variance estimate of $x(k)$ is the orthogonal projection of $x(k)$ onto the Hilbert space $L^2(Y'_k, n)$,

$$\hat{x}(k) = E[x(k)|\sigma(Y_k)] = \Pi\left[x(k)|L^2(Y'_k, n)\right]. \tag{3}$$

If the sequences $\{x(k)\}$ and $\{y(k)\}$ are jointly Gaussian, this projection is equivalent to the projection on the closed subspace $\mathcal{L}_y^k \subset L^2(Y'_k, n)$ of all affine functions of Y_k,

$$\mathcal{L}_y^k = \left\{z : \Omega \to \mathbb{R}^n : \exists T \in \mathbb{R}^{n\times l+1} : z = TY'_k\right\}. \tag{4}$$

The KF recursively computes the projection $\Pi[x(k)|\mathcal{L}_y^k]$, the best affine estimate of $x(k)$ in the minimum variance sense. This coincides with $E[x(k)|\sigma(Y_k)]$ only in the Gaussian case. When $\{x(k)\}$ and/or $\{y(k)\}$ are non-Gaussian, the computation of (3) is challenging.

Since the best affine estimate is obtained by projecting onto $\mathcal{L}_y^k$, better suboptimal estimates can be obtained by projecting the state $x(k)$ onto larger sub-spaces. For example, we may consider the space of second-order polynomial (quadratic) transformations of Y_k, denoted by $\mathcal{Q}_y^k$.

Let $Y_k^{(2)} = \mathrm{col}(Y'_k, Y_k^{[2]}) \in \mathbb{R}^{\tilde{l}}$, $\tilde{l} = 1 + l + l^2$, then

$$\mathcal{Q}_y^k = \left\{z : \Omega \to \mathbb{R}^n : \exists T \in \mathbb{R}^{n\times\tilde{l}} : z = TY_k^{(2)}\right\}. \tag{5}$$

Since $\mathcal{L}_y^k \subset \mathcal{Q}_y^k \subset L^2(Y'_k, n)$, projecting the state onto $\mathcal{Q}_y^k$ will return an estimate, having an error variance equal to or smaller than that of the affine estimate.

Theorem 1. *Suppose system (1)–(2) satisfies conditions (1)–(6). Let* $\hat{x}^{\mathcal{L}}(k) = \Pi[x(k)|\mathcal{L}_y^k]$, $\hat{x}^{\mathcal{Q}}(k) = \Pi[x(k)|\mathcal{Q}_y^k]$, *with errors* $e^{\mathcal{L}}(k) = x(k) - \hat{x}^{\mathcal{L}}(k)$, $e^{\mathcal{Q}}(k) = x(k) - \hat{x}^{\mathcal{Q}}(k)$. *Then,* $E[e^{\mathcal{L}}(k)^T e^{\mathcal{L}}(k)] \geq E[e^{\mathcal{Q}}(k)^T e^{\mathcal{Q}}(k)]$.

Proof. In virtue of the Hilbert projection theorem, $\hat{x}^{\mathcal{Q}}(k)$ has the minimum distance from $x(k)$ among all the elements of $\mathcal{Q}_y^k$. Therefore, since $\mathcal{L}_y^k \subset \mathcal{Q}_y^k$,

$$\left\|x(k) - \hat{x}^{\mathcal{L}}(k)\right\|_{L^2(X,n)}^2 \geq \left\|x(k) - \hat{x}^{\mathcal{Q}}(k)\right\|_{L^2(X,n)}^2, \tag{6}$$

and the thesis follows from the definition of the norm in $L^2(X, n)$, $\|v\|_{L^2(X,n)}^2 = \int_\Omega v^T v dP = E[v^T v]$. □

To compute the optimal quadratic estimate the idea is to derive an augmented version of (1)–(2) with vectors $\mathcal{X}(k) = \mathrm{col}(x(k), x^{[2]}(k))$, $\mathcal{Y}(k) = \mathrm{col}(y(k), y^{[2]}(k))$ and use a recursive linear filter. To this aim we have to consider the following issues:

(1) $\Pi[x(k)|\mathcal{Q}_y^k]$ must be recursively computable;
(2) the augmented system must be detectable;
(3) the noise sequences of the augmented system must be second-order asymptotically stationary processes (as defined in Carravetta et al., 1996).

As for the first point, the computation of $\Pi[x(k)|\mathcal{Q}_y^k]$ would require a growing filter size, due to the presence in $\mathcal{Q}_y^k$ of terms of the kind $y^i(k_1)y^j(k_2)$, with $i, j \leq 2$. A possible solution (De Santis et al., 1995) is to replace $Y_k^{(2)}$ with

$$\overline{Y}_k^{(2)} = \mathrm{col}\left(Y'_k, y(0)^{[2]}, \ldots, y(k)^{[2]}\right) \in \mathbb{R}^{\bar{l}}, \tag{7}$$

$\bar{l} = 1 + l + (k+1)q^2$, that is, a vector containing only the observations and their Kronecker squares from time 0 to k. We obtain the projection subspace

$$\overline{\mathcal{Q}}_y^k = \left\{z : \Omega \to \mathbb{R}^n : \exists T \in \mathbb{R}^{n\times\bar{l}} : z = T\overline{Y}_k^{(2)}\right\}, \tag{8}$$

with $\mathcal{L}_y^k \subset \overline{\mathcal{Q}}_y^k \subset \mathcal{Q}_y^k$. Thus, the corresponding *recursively computable quadratic estimate* will not be the optimal quadratic estimate, but it will still have an error variance not larger than the best linear one.

Detectability (issue 2) may not be satisfied even when the original pair (A, C) is fully observable, that is, the quadratic part of $\mathcal{X}(k)$ may not be completely observable. To solve these issues, we use output injection to rewrite (1)–(2) as an equivalent system with an asymptotically stable and hence detectable stochastic part.

This solves issue 3 as well. The problem stems from the fact that the state noise of the augmented system depends on $x(k)$, see Theorem 3.3.4 in Carravetta et al. (1996) or Theorem 1 in De Santis et al. (1995). By rewriting the system as above, the augmented system becomes asymptotically stationary.

3. The proposed approach

The proposed method, named Feedback Quadratic Filter (FQF), consists of the following steps.

(a) System (1)–(2) is rewritten as a system with a feedback given by an output injection term.
(b) The modified system is decomposed in the sum of a deterministic and a stochastic component.
(c) The augmented quadratic system is derived for the asymptotically stable stochastic component.
(d) The KF for mutually correlated state and output noises is applied to the augmented quadratic system. The final estimate is the sum of the deterministic component and the linear part of the augmented estimate.

3.1. Output injection

In the first step, the state equation (1) is transformed using the output equation (2) in

$$x(k+1) = Ax(k) + Bu(k) + f_k + Ly(k) - LCx(k) - Lg_k = \tilde{A}x(k) + Bu(k) + Ly(k) + h_k, \tag{9}$$

where $\tilde{A} = A - LC$, $L \in \mathbb{R}^{n\times q}$ is an arbitrarily chosen matrix and $h_k = f_k - Lg_k$. The following points are worth remarking.

(1) If the pair (A, C) is observable, the spectrum of $\tilde{A}$, denoted as $\rho(\tilde{A})$, can be arbitrarily assigned through L. We can therefore assume that $\rho(\tilde{A})$ is chosen in the open unit circle.
(2) The observability of (A, C) is equivalent to the observability of $(\tilde{A}, C)$, $\forall L \in \mathbb{R}^{n\times q}$, see Wonham (1979) (pag. 79, Section 3.10 as a dual result of Lemma 2.1).
(3) $\{h_k\}$ is a zero mean white sequence independent of the initial state x_0 and with finite fourth moment, by definition of $\{f_k\}$, and $\{g_k\}$.
(4) $\{h_k\}$ is correlated with the noise sequence $\{g_k\}$.

Moreover, $\psi_h^{(i)} = E\left[h_k^{[i]}\right]$, $i = 2, 3, 4$ can be computed from $\psi_f^{(i)}$ and $\psi_g^{(i)}$ as follows

$$\psi_h^{(2)} = \psi_f^{(2)} + L^{[2]}\psi_g^{(2)}, \tag{10}$$

$$\psi_h^{(3)} = \psi_f^{(3)} - L^{[3]}\psi_g^{(3)}, \tag{11}$$

$$\psi_h^{(4)} = \psi_f^{(4)} + M_2^4\left(\psi_f^{(2)} \otimes L^{[2]}\psi_g^{(2)}\right) + L^{[4]}\psi_g^{(4)}, \tag{12}$$

where M_2^4 is the coefficient matrix for the expansion of the binomial Kronecker power (see Carravetta et al., 1996).

3.2. Deterministic and stochastic components

In the second step the state sequence is split into the deterministic component and stochastic component sequences $\{x_d(k)\}$ and $\{x_s(k)\}$, respectively. The deterministic component $x_d(k)$ is the solution of

$$x_d(k+1) = \tilde{A}x_d(k) + Bu(k) + Ly(k), \qquad x_d(0) = x_0^d, \tag{13}$$

where $x_0^d = E[x_0] = \bar{x}_0$ is a deterministic variable. The stochastic component $x_s(k)$ is the solution of

$$x_s(k+1) = \tilde{A}x_s(k) + h_k, \qquad x_s(0) = x_0^s, \tag{14}$$

with $x_0^s \sim \mathcal{N}(0, \Psi_{x_0})$, and, therefore $\psi_{x_0^s}^{(i)} = E\left[x_0^{s\,[i]}\right] = \psi_{x_0}^{(i)}$. From (13) and (14) it follows that $x(k) = x_d(k) + x_s(k)$, $\forall k \geq 0$. The solution of (13) is

$$x_d(k) = \tilde{A}^k x_0^d + \sum_{\tau=0}^{k-1} \tilde{A}^{k-\tau-1}\left(Bu(\tau) + Ly(\tau)\right), \tag{15}$$

and can be computed at time k, since the initial value x_0^d and the output sequence Y_{k-1} are available. The output map (2) can be arranged as

$$y_s(k) = y(k) - Cx_d(k) = Cx_s(k) + g_k, \tag{16}$$

thus $y_s(k)$ is available at time k. We can define the corresponding sequence vector $Y_{s,k} := \mathrm{col}(y_s(j))$, $0 \leq j \leq k$, and the auxiliary vector $Y'_{s,k} = \mathrm{col}(1\ Y_{s,k})$. The vector $\overline{Y}_{s,k}^{(2)}$ and the spaces $\mathcal{L}_{y_s}^k$, $\mathcal{Q}_{y_s}^k$, $\overline{\mathcal{Q}}_{y_s}^k$ can be defined using $Y'_{s,k}$ instead of Y'_k in (7), (4), (5), and (8), respectively. Since $Y_{s,k}$ is an affine transformation of the original sequence vector Y_k and *viceversa*, see Section 4, $\mathcal{L}_y^k \equiv \mathcal{L}_{y_s}^k$. However, as we can see, in general $\overline{\mathcal{Q}}_y^k \not\equiv \overline{\mathcal{Q}}_{y_s}^k$. In the geometric approach introduced in Section 2, the recursive quadratic estimate of $x(k)$ is $\tilde{x}(k) = \Pi(x(k)|\overline{\mathcal{Q}}_{y_s}^k)$. Since the projection operator is linear,

$$\tilde{x}(k) = \Pi\left(x_d(k)|\overline{\mathcal{Q}}_{y_s}^k\right) + \Pi\left(x_s(k)|\overline{\mathcal{Q}}_{y_s}^k\right) = \tilde{x}_d(k) + \tilde{x}_s(k).$$

The projection of $x_d(k)$ onto $\overline{\mathcal{Q}}_{y_s}^k$ corresponds to itself because (15) is an affine transformation of Y_k. The estimate $\tilde{x}_s(k) = \Pi(x_s(k)|\overline{\mathcal{Q}}_{y_s}^k)$ is obtained by processing the system

$$x_s(k+1) = \tilde{A}x_s(k) + h_k, \tag{17}$$

$$y_s(k) = Cx_s(k) + g_k. \tag{18}$$

3.3. The quadratic system

By taking Kronecker squares we obtain from (17)–(18)

$$x_s^{[2]}(k+1) = \tilde{A}^{[2]}x_s^{[2]}(k) + h_k^{(2)} + \psi_h^{(2)}, \tag{19}$$

$$y_s^{[2]}(k) = C^{[2]}x_s^{[2]}(k) + g_k^{(2)} + \psi_g^{(2)}, \tag{20}$$

where

$$h_k^{(2)} = \tilde{A}x_s(k) \otimes h_k + h_k \otimes \tilde{A}x_s(k) + h_k^{[2]} - \psi_h^{(2)}, \tag{21}$$

$$g_k^{(2)} = Cx_s(k) \otimes g_k + g_k \otimes Cx_s(k) + g_k^{[2]} - \psi_g^{(2)}, \tag{22}$$

are zero-mean, temporally uncorrelated sequences, $E[h_k^{(2)}g_k^{(2)T}] \neq 0$, uncorrelated with x_0. Notice that even though f_k, g_k are stationary, the extended noises $h_k^{(2)}$, $g_k^{(2)}$ are not. This is because they are function of the state evolution, as it is apparent from (21) and (22). Let us define now the augmented state and output vectors for the stochastic system (17)–(18). By setting

$$\mathcal{A} = \begin{bmatrix} \tilde{A} & 0 \\ 0 & \tilde{A}^{[2]} \end{bmatrix}, \qquad \mathcal{C} = \begin{bmatrix} C & 0 \\ 0 & C^{[2]} \end{bmatrix}, \qquad \mathcal{H}_k = \begin{bmatrix} h_k \\ h_k^{(2)} \end{bmatrix},$$

$$\mathcal{G}_k = \begin{bmatrix} g_k \\ g_k^{(2)} \end{bmatrix}, \qquad \mathcal{U} = \begin{bmatrix} 0 \\ \psi_h^{(2)} \end{bmatrix}, \qquad \mathcal{V} = \begin{bmatrix} 0 \\ \psi_g^{(2)} \end{bmatrix},$$

the augmented system can be represented as

$$\mathcal{X}(k+1) = \mathcal{A}\mathcal{X}(k) + \mathcal{U} + \mathcal{H}_k, \tag{23}$$

$$\mathcal{Y}(k) = \mathcal{C}\mathcal{X}(k) + \mathcal{V} + \mathcal{G}_k. \tag{24}$$

The augmented noises $\{\mathcal{H}_k\}$ and $\{\mathcal{G}_k\}$ are zero-mean, temporally uncorrelated and mutually correlated sequences. We denote $Q(k) = E[\mathcal{H}_k\mathcal{H}_k^T]$ and $R(k) = E[\mathcal{G}_k\mathcal{G}_k^T]$ their covariance matrices, whereas $J(k) = E[\mathcal{H}_k\mathcal{G}_k^T]$ denotes the noise correlation matrix.

Remark 1. The eigenvalues of $\tilde{A}^{[2]}$ are the product of pairs of eigenvalues of $\tilde{A}$. Since $\rho(\tilde{A})$ is chosen in the open unit circle, matrix $\mathcal{A}$ is strictly stable. Moreover, if the eigenvalues of $\tilde{A}$ are chosen not null, $\mathcal{A}$ and consequently $[\mathcal{A}\ Q(k)]$ are FRR.

Matrix $[\mathcal{C}\ R(k)]$ might be in general not FRR, even if $[C\ \Psi_g]$ is FRR by assumption 6. This is due to redundancies introduced by the second-order Kronecker power. In order to restore this property, the augmented output is suitably reduced as described in the following. It is worth remarking that the output reduction is useful for obtaining an efficient filter implementation avoiding the processing of redundant terms, but it is not essential from the theoretical point of view.

There always exists a reduction matrix $\Sigma \in \mathbb{R}^{\bar{q}\times q^2}$, $\bar{q} < q^2$, such that $\Sigma C^{[2]}$ is FRR. By setting $y_s^{(2)}(k) := \Sigma y_s^{[2]}(k)$, (20) becomes

$$y_s^{(2)}(k) = \Sigma C^{[2]}x_s^{[2]}(k) + \Sigma g_k^{(2)} + \Sigma\psi_g^{(2)}. \tag{25}$$

Let $\mathcal{Y}_r(k) = \text{col}(y_s(k)\, {y_s}^{(2)}(k))$. By setting

$$\mathcal{C}_r = \begin{bmatrix} C & 0 \\ 0 & \Sigma C^{[2]} \end{bmatrix}, \quad \mathcal{G}_{r,k} = \begin{bmatrix} g_k \\ \Sigma g_k^{(2)} \end{bmatrix}, \quad \mathcal{V}_r = \begin{bmatrix} 0 \\ \Sigma \psi_g^{(2)} \end{bmatrix}, \quad (26)$$

the reduced version of system (23)–(24) becomes

$$\mathcal{X}(k+1) = \mathcal{A}\mathcal{X}(k) + \mathcal{U} + \mathcal{H}_k, \quad (27)$$
$$\mathcal{Y}_r(k) = \mathcal{C}_r\mathcal{X}(k) + \mathcal{V}_r + \mathcal{G}_{r,k}. \quad (28)$$

From conditions (1)–(5) in Section 2 and (10)–(12), it follows that $Q(k) = E[\mathcal{H}_k{\mathcal{H}_k}^T]$, $R_r(k) = E[\mathcal{G}_{r,k}{\mathcal{G}_{r,k}}^T]$, and $J(k) = E[\mathcal{H}_k{\mathcal{G}_{r,k}}^T]$ are given by

$$Q(k) = \begin{bmatrix} \Psi_h & st_n^{-1}\left(\psi_h^{(3)}\right) \\ \left[st_n^{-1}\left(\psi_h^{(3)}\right)\right]^T & Q_{2,2}(k) \end{bmatrix} \quad (29)$$

$$R_r(k) = \begin{bmatrix} \Psi_g & st_q^{-1}\left(\psi_g^{(3)}\right)\Sigma^T \\ \Sigma\left[st_q^{-1}\left(\psi_g^{(3)}\right)\right]^T & \Sigma R_{2,2}(k)\Sigma^T \end{bmatrix} \quad (30)$$

$$J(k) = \begin{bmatrix} -L\Psi_g & -Lst_q^{-1}\left(\psi_g^{(3)}\right)\Sigma^T \\ L^{[2]}\left[st_q^{-1}\left(\psi_g^{(3)}\right)\right]^T & J_{2,2}(k)\Sigma^T \end{bmatrix} \quad (31)$$

with

$$\begin{aligned} Q_{2,2}(k) &= \left[I_{n^2} + \mathcal{S}_{n,n}^T\right]\left\{\left[\tilde{A}\Psi_x(k)\tilde{A}^T\right] \otimes \Psi_h \right. \\ &\quad \left. + \Psi_h \otimes \left[\tilde{A}\Psi_x(k)\tilde{A}^T\right]\right\} + st_{n^2}^{-1}\left(\psi_h^{(4)}\right) - \psi_h^{(2)}{\psi_h^{(2)}}^T, \\ R_{2,2}(k) &= \left[I_{q^2} + \mathcal{S}_{q,q}^T\right]\left\{\left[C\Psi_x(k)C^T\right] \otimes \Psi_g \right. \\ &\quad \left. + \Psi_g \otimes \left[C\Psi_x(k)C^T\right]\right\} + st_{q^2}^{-1}\left(\psi_g^{(4)}\right) - \psi_g^{(2)}{\psi_g^{(2)}}^T, \\ J_{2,2}(k) &= -\left[I_{n^2} + \mathcal{S}_{n,n}^T\right]\left\{\left[\tilde{A}\Psi_x(k)C^T\right] \otimes \left(L\Psi_g\right) \right. \\ &\quad \left. + \left(L\Psi_g\right) \otimes \left[\tilde{A}\Psi_x(k)C^T\right]\right\} \\ &\quad + L^{[2]}\left[st_{q^2}^{-1}\left(\psi_g^{(4)}\right) - \psi_g^{(2)}{\psi_g^{(2)}}^T\right], \end{aligned}$$

where $\Psi_h = st_n^{-1}\left(\psi_h^{(2)}\right)$, $\Psi_x(k) = E[x_s(k){x_s}^T(k)]$, I_n is the identity in $\mathbb{R}^n$, $\mathcal{S}_{i,j}^T$ is a commutation matrix. After the output vector reduction $[\mathcal{C}_r\ R_r(k)]$ becomes

$$[\mathcal{C}_r\ R_r(k)] = \begin{bmatrix} C & 0 & \Psi_g & * \\ 0 & \Sigma C^{[2]} & * & \Sigma R_{2,2}(k)\Sigma^T \end{bmatrix}, \quad (32)$$

which is now FRR since both $[C\Psi_g]$ and $\Sigma C^{[2]}$ are FRR.

After steps (a)–(c) the reduced quadratic system assumes the form of (27)–(28). We are now in the position to show that these equations admit stationary solutions.

Theorem 2. *If L is such that the eigenvalues of $\tilde{A}$ are in the open unit circle, then $\{\mathcal{H}_k\}$, $\{\mathcal{G}_{r,k}\}$, $\mathcal{X}(k)$, and $\mathcal{Y}_r(k)$ are all second-order asymptotically stationary processes.*

Proof. In the first place $\lim_{k\to\infty}\Psi_x(k) = \Psi_x^\infty < \infty$. This follows from

$$\Psi_x(k+1) = \tilde{A}\Psi_x(k)\tilde{A}^T + \Psi_h, \quad (33)$$
$$\Psi_x(0) = E[x_s(0){x_s}^T(0)] = st_n^{-1}\left(\psi_{x_0^s}^{(2)}\right) \quad (34)$$

and the hypothesis that $\rho(\tilde{A})$ is in the open unit circle. Matrices $Q(k)$, $R_r(k)$ and $J(k)$ defined in (29)–(31) depend on $\Psi_x(k)$ and they therefore admit asymptotically stationary and bounded values $\mathcal{Q}^\infty$, R_r^∞ and J^∞ obtained by replacing $\Psi_x(k)$ with Ψ_x^∞. Since in (27)–(28) the noise processes are second-order asymptotically stationary and $\mathcal{A}$ has eigenvalues in the open unit circle, $\mathcal{X}(k)$ and $\mathcal{Y}_r(k)$ are also second-order asymptotically stationary processes. □

By using for all $k \geq 0$, $\mathcal{Q}^\infty$, R_r^∞ and J^∞ instead of $Q(k)$, $R_r(k)$ and $J(k)$ we obtain the asymptotic steady-state version of system (27)–(28).

3.4. The filtering algorithm

The KF for mutually correlated state and output noises is the linear optimal estimator for system (27)–(28) since, as mentioned, the noise sequences $\{\mathcal{H}_k\}$ and $\{\mathcal{G}_{r,k}\}$ are zero-mean, temporally uncorrelated and independent of x_0 (see De Santis et al., 1995; Liptser & Shiryayev, 2001). Moreover, as stated by Theorem 2 system (27)–(28) admits an asymptotic steady-state version. Therefore, the steady-state version of system (27)–(28) is processed by the KF for mutually correlated state and output noises (for the detailed procedure see for example Balakrishnan, 1984). This will return the projection of the extended state onto the sub-space of all affine transformations of the extended output sequence. An affine transformation of the extended outputs corresponds to a quadratic transformation of the original output, and *viceversa*. This implies that the KF will return the estimate $\tilde{\mathcal{X}}(k) = \Pi[\mathcal{X}(k)|\overline{\mathcal{Q}}_{y_s}^k]$, whose first n components are the desired recursive quadratic estimate of $x_s(k)$.

4. Properties of the feedback quadratic filter

In this section we analyze two properties of the FQF. The first one concerns the internal stability of the filter. The second one is the relationship between the estimation error and the gain of the output injection term.

Theorem 3. *If L is such that the eigenvalues of $\tilde{A}$ are in the open unit circle, then the FQF is asymptotically stationary and internally asymptotically stable.*

Proof. As stated in Section 3.4, the FQF is the KF for mutually correlated noises applied to the steady-state version of system (27)–(28). This system is asymptotically stationary (Theorem 2) and also asymptotically stable. The theorem is therefore proved since asymptotic stability implies detectability and stabilizability of the system. Indeed, for KF with or without mutually correlated noises, detectability and stabilizability are sufficient conditions to the asymptotic stationary and the internal asymptotic stability of the filter (Anderson & Moore, 1979). □

Theorem 3 motivates the proposal of the FQF since for an observable pair (A, C) it is always possible to choose L such that $\rho\tilde{A}$ is in the open unit circle. This entails the possibility of using an internally stable quadratic filter for non asymptotically stable systems.

The next theorem proves that projections on subspaces of polynomial functions of the measurements Y_k' are invariant to output injection. However, in Theorem 5 we will prove that this is not true for projections on recursively computable quadratic subspaces.

Theorem 4. *For any integer $\nu \geq 1$, let*

$$Y_k^{(\nu)} = col\left(Y_k',\ Y_k^{[2]}, \ldots, Y_k^{[\nu]}\right), \quad (35)$$
$$Y_{s,k}^{(\nu)} = col\left(Y_{s,k}',\ Y_{s,k}^{[2]}, \ldots, Y_{s,k}^{[\nu]}\right), \quad (36)$$

and $\mathcal{P}_y^k(\nu)$ and $\mathcal{P}_{y_s}^k(\nu)$ be the projection subspaces

$$\mathcal{P}_y^k(\nu) = \left\{z : \Omega \to \mathbb{R}^n : \exists T \in \mathbb{R}^{n\times\hat{l}} : z = TY_k^{(\nu)}\right\},$$
$$\mathcal{P}_{y_s}^k(\nu) = \left\{z : \Omega \to \mathbb{R}^n : \exists T \in \mathbb{R}^{n\times\hat{l}} : z = TY_{s,k}^{(\nu)}\right\}.$$

Then $\mathcal{P}_y^k(\nu) \equiv \mathcal{P}_{y_s}^k(\nu)$.

Proof. From (15), (16) it follows that, for $i = 1, \ldots, k$,

$$y_s(i) = y(i) - C\tilde{A}^i x_0^d - \sum_{\tau=0}^{i-1} C\tilde{A}^{i-\tau-1}(Bu(\tau) + Ly(\tau))$$
$$= \begin{bmatrix} \bar{y}_i & -C\tilde{A}^{i-1}L & \cdots & -CL & I_q & 0 & \cdots & 0 \end{bmatrix} Y'_k, \quad (37)$$

where $\bar{y}_i = -C\tilde{A}^i x_0^d - \sum_{\tau=0}^{i-1} C\tilde{A}^{i-\tau-1}Bu(\tau)$. This implies $Y'_k = M_k Y'_{s,k}$, with

$$M_k = \begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ \bar{y}_0 & I_q & 0 & \cdots & \cdots & 0 \\ \bar{y}_1 & -CL & I_q & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \bar{y}_k & -C\tilde{A}^{k-1}L & -C\tilde{A}^{k-2}L & \cdots & -CL & I_q \end{bmatrix}.$$

Thus, for any integer $i \geq 1$, $(Y'_k)^{[i]} = M_k^{[i]}(Y'_{s,k})^{[i]}$ from which it follows $Y_k'^{(\nu)} = \mathcal{M}_k^{(\nu)} Y_{s,k}'^{(\nu)}$ with $\mathcal{M}_k^{(\nu)} = \mathrm{diag}(M_k^{[i]})$, $i = 1, \ldots, \nu$, $Y_k'^{(\nu)}$ and $Y_{s,k}'^{(\nu)}$ defined as in (35) by using the powers of Y'_k and $Y'_{s,k}$ instead of Y_k and $Y_{s,k}$. Note that the spaces generated by the linear transformations of $Y_k'^{(\nu)}$ and $Y_{s,k}'^{(\nu)}$ coincide with $\mathcal{P}_y^k(\nu)$ and $\mathcal{P}_{y_s}^k(\nu)$, respectively. Since the matrix M_k is non-singular, lower triangular with unities in the diagonal, this property is inherited by $M_k^{[i]}$ and consequently by $\mathcal{M}_k^{(\nu)}$, thus relation $Y_k'^{(\nu)} = \mathcal{M}_k^{(\nu)} Y_{s,k}'^{(\nu)}$ is one-to-one, that implies $\mathcal{P}_y^k(\nu) \equiv \mathcal{P}_{y_s}^k(\nu)$. □

Theorem 5. *Let $\overline{\mathcal{Q}}_1^k$ and $\overline{\mathcal{Q}}_2^k$ be two versions of the space $\overline{\mathcal{Q}}_{y_s}^k$ corresponding to gain matrices L_1 and L_2. Then, in general, for any $k \geq 1$, $\overline{\mathcal{Q}}_1^k \not\equiv \overline{\mathcal{Q}}_2^k$.*

Proof. Without loss of generality, let us consider the simple case with $x_0^d = 0$ and $u(k) = 0$. Let y_1 and y_2 denote the two output vectors y_s defined in (16) by using L_1 and L_2, respectively, and $\tilde{A}_1$ and $\tilde{A}_2$ indicate the corresponding versions of $\tilde{A}$. From (15)–(16) with $\Delta L = L_2 - L_1$, we obtain

$$y_1(0) = y_2(0) = y(0),$$
$$y_1(1) = y_2(1) + C\Delta L y(0) = y_2(1) + C\Delta L y_2(0),$$
$$y_1(1)^{[2]} = y_2(1)^{[2]} + C^{[2]}\Delta L^{[2]} y_2(0)^{[2]} + M_1^2(I_q \otimes C\Delta L)(y_2(1) \otimes y_2(0)), \quad (38)$$

where M_1^2 is the coefficient matrix for the expansion of the binomial Kronecker power (see (2.2.11)–(2.2.12) in Carravetta et al. (1996)). In (38) it is clear that, if $\Delta L \neq 0$, $y_1(1)^{[2]}$ is a linear function of elements of the kind $(y_2(1) \otimes y_2(0))$ not contained in $\overline{Y}_{2,1}^{(2)}$. Thus there exists $z \in \overline{\mathcal{Q}}_1^1$ such that $z \notin \overline{\mathcal{Q}}_2^1$, which implies that $\overline{\mathcal{Q}}_1^1 \not\equiv \overline{\mathcal{Q}}_2^1$. For $k > 1$, it is easy to prove by induction that

$$y_1(k) = y_2(k) + \sum_{\tau=0}^{k-1} \Phi_\tau^k(\Delta L) y_2(\tau) \quad (39)$$

where the matrices $\Phi_\tau^k(\Delta L) \in \mathbb{R}^{q \times q}$ are, in general, different from zero when $\Delta L \neq 0$. As in the case $k = 1$, from (39) it follows that, if $\Delta L \neq 0$, $y_1^{[2]}(k)$ is a linear function of elements not contained in $\overline{Y}_k^{(2)}$. Thus there exists $z \in \overline{\mathcal{Q}}_1^k$ such that $z \notin \overline{\mathcal{Q}}_2^k$, and $\overline{\mathcal{Q}}_1^k \not\equiv \overline{\mathcal{Q}}_2^k$. □

Theorem 5 implies that the projection $\Pi(x(k)|\overline{\mathcal{Q}}_{y_s}^k)$, returned by the FQF is not invariant to the feedback gain matrix L. Therefore, the asymptotic value P_∞ of the error covariance matrix depends on L. The optimal gain L^{opt}, corresponding to the optimal recursive quadratic estimate, satisfies

$$L^{\text{opt}} = \arg\min_{|\lambda|<1, \forall\lambda\in\rho(\tilde{A})} \operatorname{trace} P_\infty(L). \quad (40)$$

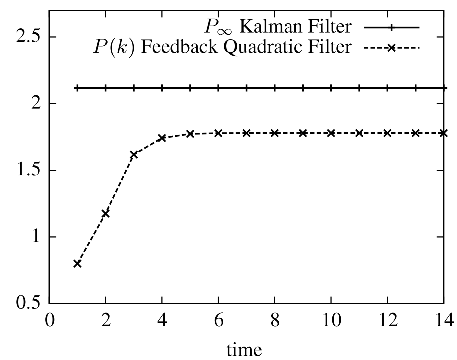


Fig. 1. Example 1, trace of the estimation error covariance matrix as a function of time for an unstable system.

Table 1
Example 1: predicted and actual estimation errors.

Filter	tr(P_∞)	MSE
Kalman Filter	2.118	2.103
Feedback Quadratic Filter	1.780	1.762

5. Numerical examples

In order to validate the proposed approach we consider three examples. In the first two examples the non-Gaussian noise sequences $\{f_k\}$ and $\{g_k\}$ are

$$f_{k_i}(\omega) = 0.4w_{f_i}(\omega) - 1.2(1 - w_{f_i}(\omega)),$$
$$g_k(\omega) = 1.5w_g(\omega) - 0.5(1 - w_g(\omega)),$$

where w_{f_i}, w_g are independent Bernoulli random variables with $p_{f_i} = 0.75$, $p_g = 0.25$.

In all the examples, the estimation accuracy of the FQF is compared with the one obtained with the standard linear KF. We remark that, since the considered systems are linear, the improvement of the estimate depends on how much non-Gaussian are the noises. If a system is well approximated by a Gaussian distribution, the improvements will be modest for any filtering algorithm, because the KF is already close to the optimum.

5.1. Example 1

Consider system (1)–(2) with:

$$A = \begin{bmatrix} 1.94 & -0.46 \\ 1.68 & 0.18 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad f_k = \begin{bmatrix} f_{k_1} \\ f_{k_2} \end{bmatrix} \in \mathbb{R}^2, \quad (41)$$

$g_k \in \mathbb{R}$, $u(k) = 0$ and initial state $x(0) = [0\ 0]^T$. Since $\rho(A) = \{1.1,\ 1.02\}$, the system is not stable. Therefore, the standard QF cannot be employed as remarked in Section 2. We want to test the behavior of the FQF and compare its performance to the KF. We choose the gain L that assigns the spectrum $\{0.05,\ 0.10\}$ to $\tilde{A}$. In Fig. 1, we plot the trace of the estimation error covariance matrix P of the FQF as a function of time k to show its fast convergence to its asymptotic value. The initial value is $P_0 = 0.4 \cdot I_2$. The trace of the stationary P_∞ matrix of the KF is plotted for comparison. The value of trace P_∞ is reduced by about 16% with respect to the KF. The reduction of the estimation error is confirmed by the mean square error (MSE) over 10^3 noises realizations of $2 \cdot 10^2$ points, with $\tilde{x}(0) = [0,\ \ 0]^T$, reported in Table 1. The good agreement between the *a priori* error and the MSE confirms the result.

5.2. Example 2

In the second example we study the dependence of the estimation error on $L$for a scalar system with $A = 0.9$, $C = 1$ and

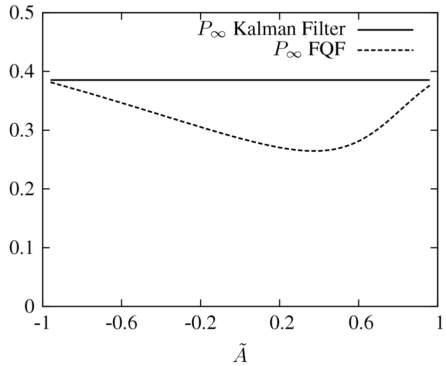


Fig. 2. Example 2: asymptotic error variance P_∞ as a function of the feedback gain.

non-Gaussian sequences f_{k_1}, g_k as above. Fig. 2 compares the value of P_∞ for the KF (which is insensitive to output feedback) and the FQF for $L \in [-0.1,\ 1.9]$, that is, $\tilde{A} \in [-1, 1]$. The figure shows that the FQF always outperforms the KF, but the error reduction varies with the gain L. The optimal value is $L^{\text{opt}} = 0.5265, \tilde{A} = \{0.3735\}$, with a reduction of P_∞ of about 31% with respect to the KF.

5.3. Example 3

We now consider the case of vector output. Consider system (1)–(2) with $n = 4, q = 2$,

$$A = \begin{bmatrix} 0.6 & 0 & 1 & 0 \\ 0 & -0.4 & 1 & 1 \\ 0 & 0 & 0.8 & 0 \\ 0 & 0 & 0 & 0.9 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \quad (42)$$

The measurement noise $g_k \in \mathbb{R}^2$ has two i.i.d. components $g_{i,k}$, $i = 1, 2$ that assume values in the interval $[-(3a+b)/4,\ (3b+a)/4]$ of width $a+b$, $a, b > 0$. The density function is $1/(2a)$ in $[-(3a+b)/4,\ (a-b)/4]$, and $1/(2b)$ in $[(a-b)/4,\ (3b+a)/4]$, thus $E(g_{i,k}) = 0$. The state noise is

$$f_k(\omega) = \begin{bmatrix} -0.95 & 0.05 \\ 2.85 & -0.15 \\ -1.8 & 0.2 \\ -2.85 & 0.15 \end{bmatrix} \begin{bmatrix} w^1 & w^2 & w^3 & w^4 \\ \bar{w}^1 & \bar{w}^2 & \bar{w}^3 & \bar{w}^4 \end{bmatrix}$$

where $w^i(\omega)$ are independent Bernoulli random variables, $p_{w_1} = p_{w_2} = p_{w_4} = 0.05$, $p_{w_3} = 0.1$ and $\bar{w}^i(\omega) = 1 - w^i(\omega)$. Since the system is stable we can compare the QF with the FQF where the gain L assigning $\rho(\tilde{A})$ inside the unit circle is chosen to minimize $\text{tr}(P_\infty)$. With the standard optimizer of MATLAB, we obtain the optimal choice of eigenvalues of $\tilde{A}$: $\lambda(\tilde{A}^{\text{opt}}) = \{0.416 \quad 0.048 \quad 0.190 \quad 0.042\}$. The size of the FQF is $n + n^2 = 20$, and no reduction is needed on the output matrix $\mathcal{C}$. We have implemented a particle filter (PF) (Arulampalam et al., 2002) suitably tuned to obtain comparable performances in terms of MSE, with $3 \cdot 10^3$ particles and systematic resampling. The result of the comparison is reported in Table 2. The MSE has been computed over 10^2 noises realizations of $5 \cdot 10^2$ points each, with $x_0 \sim \mathcal{N}(\bar{x}_0, \Psi_{x_0})$, $\Psi_{x_0} = 0.15 \cdot I_4$, $\bar{x}_0 = 0_{5\times 1}$ and $\tilde{x}(0) = \bar{x}_0$. Table 2, third column, shows that the reduction of the MSE with respect to the KF is −9.3% for the QF, −16.4% for the optimal FQF and −6.4% for the PF. The fourth column reports the average computation time for elaborating a single realization. It is evident that the computation time of the PF is definitely higher.

6. Conclusions

The FQF algorithm presented in this paper extends recursive quadratic filtering to non-Gaussian and not asymptotically stable linear systems. The proposed algorithm has the same size as the plain QF and needs the same information, namely the knowledge of the moments of the noise up to 4th order. An additional advantage is the reduction of the estimation error with respect to the plain QF. The proposed method can be readily extended to polynomial filtering (Carravetta et al., 1996).

Table 2
Example 3: comparison of estimation errors for the case of vector output.

Filter	$\text{tr}(P_\infty)$	MSE	% MSE	CPU time [ms]
KF	1.137	1.147		1.7
QF	1.030	1.040	−9.3%	47.6
FQF	0.949	0.959	−16.4%	49.7
PF	–	1.074	−6.4%	$140 \cdot 10^3$

References

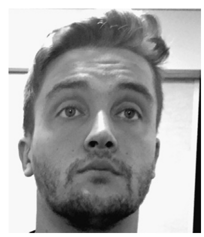
Afshar, P., Yang, F., & Wang, H. (2012). ILC-based minimum entropy filter design and implementation for non-Gaussian stochastic systems. *IEEE Transactions on Control Systems Technology*, *20*(4), 960–970.
Anderson, B. D. O., & Moore, J. B. (1979). *Optimal filtering*. Prentice-Hall, Inc..
Arasaratnam, I., Haykin, S., & Elliott, R. J. (2007). Discrete-time nonlinear filtering algorithms using Gauss–Hermite quadrature. *Proceedings of the IEEE*, *95*(5), 953–977.
Arulampalam, M. S., Maskell, S., Gordon, N., & Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, *50*(2).
Balakrishnan, A. V. (1984). *Kalman filtering theory*. New York: Optimization Software, Inc..
Bilik, I., & Tabrikian, J. (2010). MMSE-Based filtering in presence of non-Gaussian system and measurement noise. *IEEE Transactions on Aerospace and Electronic Systems*, *46*(3), 1153–1170.
Cacace, F., Conte, F., Germani, A., & Palombo, G. (2014). Quadratic Filtering for non-Gaussian and not Asymptotically Stable Linear Discrete-Time Systems. In *Proc. of the 53rd IEEE conference on decision and control CDC 14*, Los Angeles (US).
Carravetta, F., Germani, A., & Raimondi, M. (1996). Polynomial filtering for linear discrete time non-Gaussian systems. *SIAM Journal on Control and Optimization*, *34*(5), 1666–1690.
De Santis, A., Germani, A., & Raimondi, M. (1995). Optimal quadratic filtering of linear discrete-time non-Gaussian systems. *IEEE Transactions on Automatic Control*, *40*(7), 1274–1278.
Gordon, N. J., Salmond, D. J., & Smith, A. F. M. (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *Proceedings of the Institution of Electrical Engineers*, *F140*(2), 107–113.
Julier, S. J., & Uhlmann, J. K. (2004). Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, *92*(3), 401–422.
Kassam, S. A., & Thomas, J. B. (1976). Asymptotically robust detection of a known signal in contaminated non-Gaussian noise. *IEEE Transaction on Information Theory*, *22*(1), 22–26.
Liptser, R. S., & Shiryayev, A. N. (2001). *Statistics of random processes*. Berlin: Springer.
Maryak, J., Spall, J. C., & Heydon, B. (2004). Use of the Kalman filter for inference in state-space models with unknown noise distributions. *IEEE Transactions on Automatic Control*, *49*(1), 87–90.
Picinbono, B., & Devaut, P. (1988). Optimal linear-quadratic systems for detection and estimation. *IEEE Transaction on Information Theory*, *34*(2), 304–311.
Spall, J. C. (1985). Validation of state-space models from a single realization of non-Gaussian measurements. *IEEE Transactions on Automatic Control*, *30*(12), 1212–1214.
Spall, J. C. (1995). The Kantorovich inequality for error analysis of the Kalman filter with unknown noise distributions. *Automatica*, *31*(10), 1513–1517.
Spall, J. C. (2003). Estimation via Markov chain Monte Carlo. *IEEE Control Systems Magazine*, *23*(2), 34–45.
Stojanovic, V., & Nedic, N. (2015). Joint state and parameter robust estimation of stochastic nonlinear systems. *International Journal of Robust and Nonlinear Control*, http://dx.doi.org/10.1002/rnc.3490.
Stojanovic, V., & Nedic, N. (2016). Robust Kalman filtering for nonlinear multivariable stochastic systems in the presence of non-Gaussian noise. *International Journal of Robust and Nonlinear Control*, *26*, 445–460.
Verriest, E.I. (1985). Linear filters for linear systems with multiplicative noise and nonlinear filters for linear systems with non-Gaussian noise. In *IEEE American contr. conf.*.
Wonham, W. M. (1979). *Linear multivariate control: a geometric approach*. New York: Springer-Verlag, Inc..
Wu, H., & Chen, G. (1993). Suboptimal Kalman filtering for linear systems with non-Gaussian noise. *Approximate Kalman Filtering*, *2*, 113–136.
Zakai, M. (1969). On the optimal filtering of diffusion processes. *Probability Theory and Related Fields*, *11*((3)3), 230–243.

Zhang, J., Kuai, Y., Ren, M., Luo, Z., & Lin, M. (2016). Event-triggered distributed filtering for non-Gaussian systems over wireless sensor networks using survival information potential criterion. *IET Control Theory & Applications*, *10*(13), 1524–1530.

Filippo Cacace graduated in electronic engineering at Politecnico di Milano in 1988, where he received the Ph.D. Degree in computer science in 1992. Since 2003 he has been working at the Faculty of Engineering of the University Campus Bio-Medico of Rome, where he is currently an Associate Professor. His current research interests include nonlinear systems and observers, stochastic and delay systems, system identification and applications to systems biology.

Francesco Conte was born in Popoli (PE), Italy, in 1984. He received the Master Degree in computer science and automatic engineering and the Ph.D. Degree in electrical and information engineering both from the University of L'Aquila, Italy, in 2009 and 2013, respectively. Currently, he is a Postdoc with the Department of Electrical, Electronics and Telecommunication Engineering and Naval Architecture (DITEN), University of Genoa, Genova, Italy. From December 2008 to April 2009, he was a visiting scholar with the French National Institute for Research in Computer Science and Control (INRIA), Grenoble, France. He is author of more than 30 scientific publications. His research interests include nonlinear and non-Gaussian systems filtering, analysis, estimation and control of delay and nonlinear systems, power system modeling and control with a particular focus on the smart control of distributed generation sources, controllable loads and energy storage systems.

Alfredo Germani received the Laurea Degree in physics and the Post-doctoral Degree in computer and system engineering from University of Rome La Sapienza, Rome, Italy, in 1972 and 1974, respectively. From 1975 to 1986, he was a Researcher at Istituto di Analisi dei Sistemi e Informatica A. Ruberti of the Italia National Research Council, Rome. In 1978 and 1979, he was a Visiting Scholar with the Department of System Science, University of California at Los Angeles. From 1986 to 1987, he was a Professor of Automatic Control at the University of Calabria, Italy. Since 1987 he is Full Professor of System Theory at Universit of L'Aquila, Italy where from 1989 to 1992 he was the Chairperson of the School of Electronic Engineering. At University of L'Aquila, from 1995 to 2000 he led the Department of Electrical and Computer Sciences Engineering, and from 1996 to 2009 he has been the Coordinator of the Ph.D. School in Electronics. From 2007 to 2010 he has been a member of the Advisory Board of the Engineering School. Since 2012 he has been a member with the Department of Information Engineering, Computer Sciences and Mathematics. He has published more than 160 research papers in books, international journals and refereed conference proceedings in the fields of systems theory, systems identification and data analysis, nonlinear, stochastic and optimal control theory, distributed and delay systems, finitely additive white noise theory, approximation theory, optimal polynomial filtering for non-Gaussian systems, image processing and restoring, and mathematical modeling for biological processes.



Giovanni Palombo received the Master Degree in Biomedical Engineering in 2012 and the Ph.D. Degree in Automatics and System Theory in 2016, respectively from Università Campus Bio-Medico di Roma and Università degli Studi dell' Aquila. His current research interests focus mainly on non-Gaussian filtering, delay systems and SDE.