Retrodictor-corrector filter

Published Version:
Modenini, D., Zannoni, M., Tortora, P. (2017). Retrodictor-corrector filter. JOURNAL OF GUIDANCE CONTROL AND DYNAMICS, 40(9), 2329-2933 [10.2514/1.G002969].

Availability:
This version is available at: https://hdl.handle.net/11585/622488 since: 2018-02-19

Published:

DOI: http://doi.org/10.2514/1.G002969

A Retrodictor-Corrector Filter

Dario Modenini,[1] Marco Zannoni,[2] and Paolo Tortora[3]
Università di Bologna, 47121 Forlì, Italy

I. Introduction

Nonlinear filtering of dynamical systems is an extensively studied problem, usually tackled extending the well-established optimality results existing for linear systems, i.e. the Kalman filter, in the realm of nonlinear functions. The approaches are numerous [1], starting from the early developed linearized Kalman filter and extended Kalman filter (EKF), soon refined to attain better performance in case of severe nonlinearities. Some methods retain higher order derivatives as with the second order EKF, others rely on iterative linearization. This last concept was most often applied to the measurement equation, leading to the iterated extended Kalman filter (IEKF), which improves upon the EKF by iterating the nonlinear measurement update equation through a re-linearization about the updated state estimate, while retaining the time update step equal to the EKF one. The application of the re-linearization technique to handle also nonlinear dynamics received apparently less attention. An early example is represented by a single stage iterated filter/smoother (SSIFS), proposed by Wishner et al. in [2], which embeds a smoothing step within the filter to improve the past estimate when a new measurement is processed. The smoothed estimate is then used as a starting point to re-linearize the prediction step for the filtering stage. This concept was more recently generalized and merged with a moving window batch filter by Psiaki with his backward smoothing extended Kalman filter (BSEKF) [3], allowing the smoothing process to run into the past for more than one stage. Such a combined filter-smoother was shown to outperform both the EKF and the unscented Kalman filter (UKF) in a spacecraft attitude estimation problem. This

[1] Research Fellow, Department of Industrial Engineering (DIN), Via Fontanelle 40; dario.modenini@unibo.it.(Corresponding Author)
[2] Research Fellow, Department of Industrial Engineering (DIN), Via Fontanelle 40; m.zannoni@unibo.it.
[3] Associate Professor, Department of Industrial Engineering (DIN), Via Fontanelle 40, AIAA Senior Member; paolo.tortora@unibo.it.

was achieved at the expense of retaining at least 30 stages of smoothing window, resulting in more than 100 times the computational burden of a standard EKF, thus confining such a filter mainly to off-line applications.

As an alternative to the above methods, a quite different, yet very effective, approach was the one developed through the so-called Sigma-points filters [4], based on the concept of *statistical linearization* [1], the most popular of which is probably the UKF [5]. Among the advantages of the Sigma-points filters, is the absence of Jacobian matrices to be computed.

The algorithm investigated in this Note, which we call the retrodictor-corrector filter (RCF), belongs to the first family of filters and consists of a single-stage recursive scheme, based upon a generalization of the iterated extended Kalman filter. The motivating idea stems from recognizing an approximation shared by the nonlinear recursive methods cited above, which is inherent to their *predictor-corrector* structure. Having obtained estimates for the state and error covariance matrix at a certain time instant, say $k-1$, the estimate at the subsequent measurement instant k is computed through two steps: first, the forward time propagation of both the state vector and error covariance matrix (*predictor step*) and then by incorporating the information coming from the measurements (*corrector step*). Despite its intuitive nature, the predictor step can be performed exactly only in the linear filtering problem, leading to the well-celebrated KF, while its extension to the nonlinear case gives rise to a certain degree of approximation, due to the fact that one needs to propagate through a nonlinear transformation the first and second moment of a random variable.

In the attempt of mitigating the problems associated with the nonlinear prediction step, we will take the perspective of recursive filtering as the result of cost function minimization, following a wide literature on this (e.g. [6, pp. 361–482], [7]), with the cost function reflecting the balance between the measurement residuals and the so-called a-priori residuals. These last are usually built upon the one-step ahead prediction of the estimate obtained at the previous measurement time, and weighted using the inverse predicted error covariance. Here instead, we will employ as a source of a-priori information the previous estimate directly, to which the current state shall be retrofitted. With such formulation, a novel concept is brought forward, named the backward equivalent process noise. Its role is to account for the uncertainty in the retrofit of the current estimate due to the stochastic nature of the governing differential equation, as will be thoroughly discussed in Section II. The nonlinear minimization problem which leads to the desired estimate will be tackled using an iterative differential correction scheme.

In this work, we address first the basic mathematical specification of the retrodictor-corrector filter. Then a solution method is shown, which allows to highlight several similarities with already existing filters. The performance of the algorithm is tested through numerical simulations involving the gyroless angular velocity estimation problem for a tumbling spacecraft.

II. Basic specification of the nonlinear estimator

We consider the problem of a dynamic state vector evolving in time according to the nonlinear differential equation:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}, t) + \boldsymbol{G}(t)\boldsymbol{w}(t) \tag{1}$$

where $\boldsymbol{f}$ is an m-dimensional nonlinear functions of the state vector $\boldsymbol{x}(t)$, $\boldsymbol{G}$ is a $m \times p$ matrix whose entries are in general time dependent, and $\boldsymbol{w}(t)$ is a p-dimensional zero-mean, white random process with known autocovariance $E[\boldsymbol{w}(t)\boldsymbol{w}^T(t-\tau)] = \boldsymbol{Q}(t)\delta(\tau)$. We further assume that the observation $\boldsymbol{z}_k$ with dimension n approximately tracks this process as in:

$$\boldsymbol{z}_k = \boldsymbol{h}_k(\boldsymbol{x}_k) + \boldsymbol{v}_k \tag{2}$$

where $\boldsymbol{h}_k$ is the nonlinear observation function and $\boldsymbol{v}_k$ is a zero-mean, white, random noise vector of known covariance matrix $\boldsymbol{R}_k$.

Given the time series of data $\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_n$, an a-priori state vector estimate $\tilde{\boldsymbol{x}}_0$ and associated error covariance matrix $\boldsymbol{P}_0$, we generate the sequence of optimal state estimates $\tilde{\boldsymbol{x}}_k, k = 1, 2, ...n$ by recursively minimizing a cost functional defined as:

$$\tilde{\epsilon}_k \triangleq \|\tilde{\boldsymbol{x}}_{k-1} - \tilde{\boldsymbol{x}}_{k-1,k}\|^2_{\boldsymbol{P}^{-1}_{k-1,k}} + \|\boldsymbol{z}_k - \boldsymbol{h}_k(\tilde{\boldsymbol{x}}_k)\|^2_{\boldsymbol{R}^{-1}_k} \tag{3}$$

In Eq.(3), $\tilde{\boldsymbol{x}}_{k-1,k}$ and $\boldsymbol{P}_{k-1,k}$ are, respectively, the retrodicted estimate and error covariance at t_{k-1} starting from t_k [15]. The former is usually computed through numerical integration of the noise free version of Eq.(1) starting from the current estimate [16], unless an exact solution is available. To define $\boldsymbol{P}_{k-1,k}$, we first recognize that we are considering the previous estimate, $\tilde{\boldsymbol{x}}_{k-1}$, as an additional measurement, and the system dynamic, integrated backward in time, as the corresponding measurement function. Such additional measurement is corrupted by a noise vector, $\boldsymbol{e}_{k-1,k} \triangleq \tilde{\boldsymbol{x}}_{k-1} - \boldsymbol{x}_{k-1,k}$, which differs from the estimation error at the previous time step, $\boldsymbol{e}_{k-1}$, because of the random process affecting the system dynamic.

Indeed, we may formally define a discrete noise, $\boldsymbol{u}_{k-1,k}$, which maps the effects of the continuous random process $\boldsymbol{w}(t)$ backward in time from instant k to $k-1$, through integration of Eq.(1):

$$\boldsymbol{x}_{k-1} = \boldsymbol{x}_k + \int_{t_k}^{t_{k-1}} \boldsymbol{f}(\boldsymbol{x},\tau)d\tau + \int_{t_k}^{t_{k-1}} \boldsymbol{G}(\tau)\boldsymbol{w}(\tau)d\tau = \boldsymbol{x}_{k-1,k} + \boldsymbol{u}_{k-1,k} \tag{4}$$

That is:

$$\boldsymbol{u}_{k-1,k} = \int_{t_k}^{t_{k-1}} \boldsymbol{G}(\tau)\boldsymbol{w}(\tau)d\tau \tag{5}$$

which is a zero-mean Wiener process. Its covariance is:

$$\boldsymbol{Q}_{k-1,k} = \int_{t_k}^{t_{k-1}} \boldsymbol{G}(\tau)\boldsymbol{Q}(\tau)\boldsymbol{G}^T(\tau)d\tau \tag{6}$$

whose evaluation reduces to a quadrature problem, which can be computed with desired accuracy. Then, by substituting $\boldsymbol{x}_{k-1,k}$ from Eq.(4) into the definition of $\boldsymbol{e}_{k-1,k}$, one retrieves:

$$\boldsymbol{e}_{k-1,k} = \boldsymbol{e}_{k-1} + \boldsymbol{u}_{k-1,k} \tag{7}$$

The net effect of the backward process noise is thus an inflation of the effective noise when trying to retrofit the current estimate to the previous one, similarly as what happens with the KF projection ahead step. The a-priori covariance matrix to be employed in the cost function follows as:

$$\boldsymbol{P}_{k-1,k} \triangleq E[\boldsymbol{e}_{k-1,k}\boldsymbol{e}_{k-1,k}^T] = E[\boldsymbol{e}_{k-1}\boldsymbol{e}_{k-1}^T] + E[\boldsymbol{u}_{k-1,k}\boldsymbol{u}_{k-1,k}^T] = \boldsymbol{P}_{k-1} + \boldsymbol{Q}_{k-1,k} \tag{8}$$

which holds under the assumption of $\boldsymbol{e}_{k-1}$ uncorrelated with $\boldsymbol{u}_{k-1,k}$.

Note that $\boldsymbol{P}_{k-1,k}$ can be regarded as the true covariance of $\boldsymbol{e}_{k-1,k}$ noise, and for its computation we did not resort to any approximation: nonlinear transformations of the state error are not involved in the derivation, as it happens when projecting ahead the covariance matrix in a nonlinear prediction step.

III. Solution to the nonlinear minimization problem

Suppose that a nominal guess for $\tilde{\boldsymbol{x}}_k$, denoted with $\bar{\boldsymbol{x}}_k$, is available. Then a solution to the minimization of Eq.(3) can be obtained with the iterated Gauss-Newton method [8], in terms of the residuals:

$$\begin{aligned} \delta\tilde{\boldsymbol{x}}_{k-1} &= \tilde{\boldsymbol{x}}_{k-1} - \bar{\boldsymbol{x}}_{k-1,k} \\ \delta\boldsymbol{z}_k &= \boldsymbol{z}_k - \boldsymbol{h}_k(\bar{\boldsymbol{x}}_k) \end{aligned} \tag{9}$$

and of the their partial derivatives with respect to the unknown state:

$$\begin{aligned} \boldsymbol{H}_k &= \left.\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}\right|_{\bar{\boldsymbol{x}}_k} \\ \boldsymbol{\Phi}_{k-1,k} &= \left.\frac{\partial \boldsymbol{x}_{k-1,k}}{\partial \boldsymbol{x}}\right|_{\bar{\boldsymbol{x}}_k} \end{aligned} \tag{10}$$

The state transition matrix, $\boldsymbol{\Phi}$, can also be obtained by integrating the linearized state dynamics according to:

$$\frac{\mathrm{d}\boldsymbol{\Phi}(t,\tau)}{\mathrm{d}t} = \boldsymbol{F}(t)\boldsymbol{\Phi}(t,\tau) \qquad \boldsymbol{\Phi}(\tau,\tau) = \boldsymbol{I} \tag{11}$$

with $\boldsymbol{F}(t) = \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right|_{\bar{\boldsymbol{x}}(t)}$.

A Gauss-Newton step, $\delta\tilde{\boldsymbol{x}}_k$, towards the minimization of Eq.(3) reads:

$$\begin{aligned} \delta\tilde{\boldsymbol{x}}_k = &\left[\boldsymbol{\Phi}_{k-1,k}^T \boldsymbol{P}_{k-1,k}^{-1} \boldsymbol{\Phi}_{k-1,k} + \boldsymbol{H}_k^T \boldsymbol{R}_k^{-1} \boldsymbol{H}_k\right]^{-1} \\ &\left[\boldsymbol{\Phi}_{k-1,k}^T \boldsymbol{P}_{k-1,k}^{-1} \delta\tilde{\boldsymbol{x}}_{k-1} + \boldsymbol{H}_k^T \boldsymbol{R}_k^{-1} \delta\boldsymbol{z}_k\right] \end{aligned} \tag{12}$$

A more familiar form of the solution step above can be retrieved as follows. First we define the matrix $\boldsymbol{P}_{k,k-1}$ by forward mapping of $\boldsymbol{P}_{k-1,k}$:

$$\boldsymbol{P}_{k,k-1} \triangleq \boldsymbol{\Phi}_{k,k-1} \boldsymbol{P}_{k-1,k} \boldsymbol{\Phi}_{k,k-1}^T \tag{13}$$

which closely resembles the a-priori covariance matrix $\boldsymbol{P}_{k|k-1}$ from (E)KF. Then we let:

$$\boldsymbol{P}_k^{-1} = \boldsymbol{\Phi}_{k-1,k}^T \boldsymbol{P}_{k-1,k}^{-1} \boldsymbol{\Phi}_{k-1,k} + \boldsymbol{H}_k^T \boldsymbol{R}_k^{-1} \boldsymbol{H}_k = \boldsymbol{P}_{k,k-1}^{-1} + \boldsymbol{H}_k^T \boldsymbol{R}_k^{-1} \boldsymbol{H}_k \tag{14}$$

Substituting into Eq.(12) yields to:

$$\delta\tilde{\boldsymbol{x}}_k = \boldsymbol{P}_k \left[\boldsymbol{P}_{k,k-1}^{-1} \boldsymbol{\Phi}_{k,k-1} \delta\tilde{\boldsymbol{x}}_{k-1} + \boldsymbol{H}_k^T \boldsymbol{R}_k^{-1} \delta\boldsymbol{z}_k\right] \tag{15}$$

Then, by adding and subtracting the product $\boldsymbol{H}_k^T \boldsymbol{R}_k^{-1} \boldsymbol{H}_k \boldsymbol{\Phi}_{k,k-1} \delta\tilde{\boldsymbol{x}}_{k-1}$ to the right hand side of Eq.(15) and by making use of Eq.(14), we obtain:

$$\delta\tilde{\boldsymbol{x}}_k^{(n+1)} = \boldsymbol{\Phi}_{k,k-1} \delta\tilde{\boldsymbol{x}}_{k-1} + \boldsymbol{K}_k \left[\delta\boldsymbol{z}_k - \boldsymbol{H}_k \boldsymbol{\Phi}_{k,k-1} \delta\tilde{\boldsymbol{x}}_{k-1}\right] \tag{16}$$

where we have introduced the gain $\boldsymbol{K}_k = \boldsymbol{P}_k \boldsymbol{H}_k^T \boldsymbol{R}_k^{-1}$.

This final form highlights the retrodictor-corrector structure of the proposed estimator: the retrodicted state vector residual $\delta\tilde{\boldsymbol{x}}_{k-1}$ is mapped at current time and then corrected with the new available data through the innovation term $[\delta\boldsymbol{z}_k - \boldsymbol{H}_k \boldsymbol{\Phi}_{k,k-1} \delta\tilde{\boldsymbol{x}}_{k-1}]$.

A reasonable choice to start the iteration is by forward propagation of the previous estimate, i.e. by setting $\bar{\boldsymbol{x}}_k^{(0)} = \tilde{\boldsymbol{x}}_{k|k-1}$. This way, $\delta\tilde{\boldsymbol{x}}_{k-1}$ is null, and the update equation reduces to:

$$\tilde{\boldsymbol{x}}_k = \tilde{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k \left[\boldsymbol{z}_k - \boldsymbol{h}_k(\tilde{\boldsymbol{x}}_{k|k-1})\right] \tag{17}$$

which is the popular EKF update formula.

The iterative solution scheme can be cycled by setting: $\bar{\boldsymbol{x}}_k^{(n+1)} = \bar{\boldsymbol{x}}_k^{(n)} + \delta\tilde{\boldsymbol{x}}_k^{(n)}$, re-evaluating the residuals and partial derivatives, Eqs.(9)-(10). Note that, in a standard IEKF, only the measurement matrix $\boldsymbol{H}_k$ is updated along iterations, while here the transition matrix as well. The iterations should be stopped upon convergence, i.e. whenever the relative decrease in the cost function ceases to be significant: then, the computational burden required to run the entire algorithm will be comparable to the one of an EKF times the number of iterations performed.

IV. Reduction to limiting cases

We will now show some properties of the proposed scheme emerging under some simplifications or limiting cases, to highlight its connections with other classic algorithms.

A. The linear case

First we consider the fully linear estimation problem: the filter specification runs in parallel, upon substitution of the functions $(\boldsymbol{f}, \boldsymbol{G})$ with the state and process noise transition matrices $(\boldsymbol{\Phi}_k, \boldsymbol{\Gamma}_k)$ and of the function $\boldsymbol{h}$ with the observation matrix $\boldsymbol{H}_k$. Then the cost function to be minimized reads:

$$\tilde{\epsilon}_k = \|\tilde{\boldsymbol{x}}_{k-1} - \boldsymbol{\Phi}_{k-1,k}\tilde{\boldsymbol{x}}_k\|^2_{\boldsymbol{P}^{-1}_{k-1,k}} + \|\boldsymbol{z}_k - \boldsymbol{H}_k\tilde{\boldsymbol{x}}_k\|^2_{\boldsymbol{R}_k^{-1}} \tag{18}$$

which shall be compared with the cost function minimized by the linear KF [6, pp. 466–470]:

$$\tilde{\epsilon}_k = \|\boldsymbol{\Phi}_{k,k-1}\tilde{\boldsymbol{x}}_{k-1} - \tilde{\boldsymbol{x}}_k\|^2_{\boldsymbol{P}^{-1}_{k|k-1}} + \|\boldsymbol{z}_k - \boldsymbol{H}_k\tilde{\boldsymbol{x}}_k\|^2_{\boldsymbol{R}_k^{-1}} \tag{19}$$

The a-priori covariance $\boldsymbol{P}_{k|k-1}^{-1}$ in the linear case is equivalent to $\boldsymbol{P}_{k,k-1}^{-1}$ as defined in Eq.(13) (see Appendix). It is then quite straightforward to show that the first terms at the right hand sides of Eqs. (18) and (19) (i.e. the a-priori residuals) are actually the same. That is, in the linear case only, the prediction and the retrodiction perspectives are equivalent, thus showing complete consistency of the least square criterion proposed in Eq.(3) with KF.

B. Relations to the Iterated Extended Kalman Filter

Next we may inspect the relation existing between the RCF and the IEKF. The cost functions minimized by the IEKF reads [6, pp. 440–442]:

$$\tilde{\epsilon}_k = \left\|\tilde{\boldsymbol{x}}_k - \tilde{\boldsymbol{x}}_{k|k-1}\right\|^2_{\boldsymbol{P}_{k|k-1}^{-1}} + \|\boldsymbol{z}_k - \boldsymbol{h}_k(\tilde{\boldsymbol{x}}_k)\|^2_{\boldsymbol{R}_k^{-1}} \tag{20}$$

which shall be compared to Eq.(3).

The solutions in terms of Gauss-Newton iterations are given by Eq.(15) and, for the IEKF as:

$$\delta\tilde{\boldsymbol{x}}_k^{(n+1)} = \boldsymbol{P}_k \left[\boldsymbol{P}_{k|k-1}^{-1}\delta\tilde{\boldsymbol{x}}_k + \boldsymbol{H}_k^T \boldsymbol{R}_k^{-1}\delta\boldsymbol{z}_k\right] \tag{21}$$

The difference between the IEKF and the RCF relies in the different definition of the a-priori residuals in Eqs.(20) and (3): the IEKF employs an a-priori estimate which is derived through a forward prediction of the mean and covariance of the previous estimate. This difference is reflected also in the solution scheme: in Eq.(21) the a-priori covariance matrix $\boldsymbol{P}_{k|k-1}$ is not updated through the iteration cycles, since the transition matrix used for the forward mapping is assumed constant, being obtained from a *past* estimate. Conversely, for the proposed filter $\boldsymbol{P}_{k,k-1}$ is recomputed at each cycle, being mapped using a transition matrix which is obtained from the *current* estimate. It can be further shown that the outcome of Eq.(15) and Eq.(21) are equivalent at the first iteration only, when both are initialized with the forward map of $\tilde{\boldsymbol{x}}_{k-1}$. In that situation, as previously noted, $\boldsymbol{P}_{k,k-1}$ is equivalent to the a-priori covariance of the standard EKF formulation.

A similar argument holds also for the UKF, whose measurement update equation is derived according to the Bayesian linear minimum mean-squared error estimator [5]. It makes use of the predicted mean $\tilde{\boldsymbol{x}}_{k|k-1}$ and predicted error covariance $\boldsymbol{P}_{k|k-1}$ as well, but differs from the EKF in the method used to generate such predictions, which typically achieves better accuracy. This, however, does not eliminate the approximation intrinsic in the nonlinear predictor step.

C. The Deterministic Case

Up to know we have considered systems whose dynamic is driven by a white random process. Nevertheless, KF proved to work effectively also in many practical filtering applications where the unknown external disturbances and/or model uncertainties cannot be regarded as random processes. In such cases, one may rather conceive the estimation problem in a purely deterministic setting [6, pp. 303–314] and try to fit the current estimate to all the available data, as in:

$$\epsilon_k(\tilde{\boldsymbol{x}}_k) = \sum_{h=0}^{k} \left(\boldsymbol{z}_{k-h} - \boldsymbol{h}_{k-h}(\tilde{\boldsymbol{x}}_{k-h,k})\right)^T \boldsymbol{R}_{k-h}^{-1} \left(\boldsymbol{z}_{k-h} - \boldsymbol{h}_{k-h}(\tilde{\boldsymbol{x}}_{k-h,k})\right) \tag{22}$$

In a deterministic framework, compensation of modeling errors is achieved by reducing the filter memory length and/or applying a fading factor which progressively de-weights older residuals in the build-up of the cost function [9]. Such a technique was proven to work effectively in a linear height estimation problem [10].

Eq.(22) can be rewritten as a one-step recursion:

$$\epsilon_k(\tilde{\boldsymbol{x}}_k) = \epsilon_{k-1}^*(\tilde{\boldsymbol{x}}_k) + \left(\boldsymbol{z}_k - \boldsymbol{h}_k(\tilde{\boldsymbol{x}}_k)\right)^T \boldsymbol{R}_k^{-1} \left(\boldsymbol{z}_k - \boldsymbol{h}_k(\tilde{\boldsymbol{x}}_k)\right) \tag{23}$$

Note that $\epsilon_{k-1}^*(\tilde{\boldsymbol{x}}_k) \neq \epsilon_{k-1}(\tilde{\boldsymbol{x}}_{k-1})$. It is now easy to show that the retrodictor-corrector filtering scheme with $\boldsymbol{Q}(t) = \boldsymbol{0}$ provides an approximate solution to the minimization of Eq.(22). First we recognize that in the deterministic case $\boldsymbol{P}_{k-1,k}$ reduces to $\boldsymbol{P}_{k-1}$. Then, the first term in Eq.(3) (i.e. the a-priori residual) can be seen as a Taylor expansion of $\epsilon_{k-1}^*(\tilde{\boldsymbol{x}}_k)$ around $\epsilon_{k-1}(\tilde{\boldsymbol{x}}_{k-1})$ [3] which, up to second order, leads to:

$$\epsilon_{k-1}^*(\tilde{\boldsymbol{x}}_k) \approx \left(\tilde{\boldsymbol{x}}_{k-1} - \tilde{\boldsymbol{x}}_{k-1,k}\right)^T \boldsymbol{P}_{k-1}^{-1} \left(\tilde{\boldsymbol{x}}_{k-1} - \tilde{\boldsymbol{x}}_{k-1,k}\right) + \text{const.} \tag{24}$$

In the equation above the first order term is not present since the gradient of the cost function at minimization is null and the $\approx$ sign is due to the fact that $\boldsymbol{P}_{k-1}^{-1}$ is only an approximation of the Hessian of $\epsilon_{k-1}(\tilde{\boldsymbol{x}}_{k-1})$.

V. Numerical Results: gyroless rate estimation for a tumbling spacecraft

In this Section filter validation is addressed by implementing the proposed algorithm within a simulated test case. The proposed RCF will be compared to the EKF and the UKF. This latter requires three tunable parameters to control the spread of the sigma-points used for the statistical linearization: in this work we used values which are considered standard good first guesses [17], without seeking for complicated multidimensional tuning procedures. This is one of the reasons why we will not make any absolute claim about the relative merits between our filter and the UKF.

We consider the method presented in [11] for fast estimation of the angular rate of a tumbling spacecraft from sequential readings of the Earth's magnetic field. This same application was also used in [12] to explore the self-tuning properties of a recursive fading memory filter. Here we use this same test case since it easily allows to investigate the effects of a variable "degree of nonlinearity" of the system dynamic on the filters' performances, by changing the magnitude of the initial angular rate vector. If the external disturbance torques are approximated as a zero mean stationary process, $\zeta(\mathrm{t})$, the vector stochastic differential equation representing the dynamic model is given by:

$$\boldsymbol{J}\dot{\omega} = \omega \times \boldsymbol{J}\omega + \zeta \tag{25}$$

where ω is the SC angular rate vector and $\boldsymbol{J}$ is the matrix of inertia. The observation model is based on the equation

$$\frac{\mathrm{d}\boldsymbol{b}}{\mathrm{d}t} - \frac{\partial \boldsymbol{b}}{\partial t} = \boldsymbol{\omega} \times \boldsymbol{b} \tag{26}$$

where $\boldsymbol{\omega}$ is the spacecraft angular rate vector, $\boldsymbol{b}$ is the Earth's magnetic field vector, $\partial \boldsymbol{b}/\partial t$ is the temporal derivative of the magnetic field vector, taken in the body-fixed frame, and $d\boldsymbol{b}/dt$ is the (total) temporal derivative taken in an inertial reference frame. During detumbling, and for most orbits, this last is negligible, which yields

$$\frac{\partial \boldsymbol{b}}{\partial t} \approx -\boldsymbol{\omega} \times \boldsymbol{b} \equiv [\boldsymbol{b}\times]\, \boldsymbol{\omega} \tag{27}$$

where $[\boldsymbol{b}\times]$ is the cross-product matrix corresponding to the magnetic field vector.

The Three-Axis Magnetometer (TAM) reading at time t_k is related to the true magnetic field via $\hat{\boldsymbol{b}}_k = \boldsymbol{b}_k + \boldsymbol{n}_k$, where the TAM stationary measurement noise is distributed as $\boldsymbol{n}_k \sim N\left(\boldsymbol{0}, \boldsymbol{R}_{\mathbf{TAM}}\right)$, with known covariance $\boldsymbol{R}_{TAM}$. The body-referenced temporal derivative is approximated using a first-order backward finite difference, computed using two successive TAM readings, assumed sampled at 50 Hz. The observation equation is then written as

$$\boldsymbol{z}_k = \boldsymbol{H}_k\, \boldsymbol{x}_k + \boldsymbol{v}_k \tag{28}$$

where $\boldsymbol{H}_k = \left[\hat{\boldsymbol{b}}_k \times\right] \Delta t$ is the time varying observation matrix, $\boldsymbol{z}_k = \hat{\boldsymbol{b}}_k - \hat{\boldsymbol{b}}_{k-1}$ is the effective measurement vector, and $\boldsymbol{v}_k$ is the effective measurement noise.

The differencing of TAM readings, which is performed to generate the filter measurement $\boldsymbol{z}_k$, makes the effective measurement noise colored and state-dependent. For the sake of simplicity, the procedure proposed in [11] to handle colored noise is not implemented here. Rather, the synthetic measurements fed to the filters are obtained in simulation by differencing punctual magnetometer readings; the outcome is then corrupted by an additive zero-mean white Gaussian noise.

The performance of the three algorithms are compared through simulation, for a spacecraft traveling on a circular orbit with a radius of 6 905 km, an inclination of 97 deg, and a period of 5 710 s. The real attitude motion is numerically integrated in the presence of a driving process noise with power spectral density equal to $\boldsymbol{Q} = 10^{-14} \cdot \boldsymbol{I}_3$ N^{2}m^2, and a diagonal inertia matrix with entries [2.1, 2.05, 1.5] kgm^2. The simulated noise-free TAM readings are generated from a tenth-order International Geomagnetic Reference Field model. Once differenced, a zero-mean white Gaussian noise with covariance $\boldsymbol{R}_{MAG} = \boldsymbol{I}_3 \cdot 10^{-13}$ T^2 is added. The filter is run at 0.5 Hz, with initial estimate set to $\tilde{\boldsymbol{x}}_0 = [0,\ 0,\ 0]^T$ rad/s and the covariance matrix is initialized with $\boldsymbol{P}_0 = \boldsymbol{I}_3$ rad^2/s^2.

The performance of the predictor-corrector algorithm are compared with the EKF and UKF by running an ensemble of 100 Monte Carlo simulations, with the S/C initial angular rate vector having a random direction given a fixed magnitude. Three set of simulations were performed, for angular rate magnitudes of 20 deg/s, 30 deg/s, and 40 deg/s: the corresponding root mean square (rms) error histories, averaged over the Monte Carlo runs, are shown in Figures 1, 2 and 3. These highlight the benefits of the retrodictor-corrector formulation for handling the nonlinear system dynamics. In particular, Figure 1 shows that for the lower angular rate magnitude, the three filters provide practically the same accuracy. However, when the initial rate raises, the EKF performs worse both during the initial transient and at steady state, with a difference becoming increasingly evident at $|\omega| = 40$ deg/s (see Figure 3), while the RCF and the UKF remains on par. Note that tumbling rates of such magnitude, though not common, have been encountered during actual missions [13], [14].

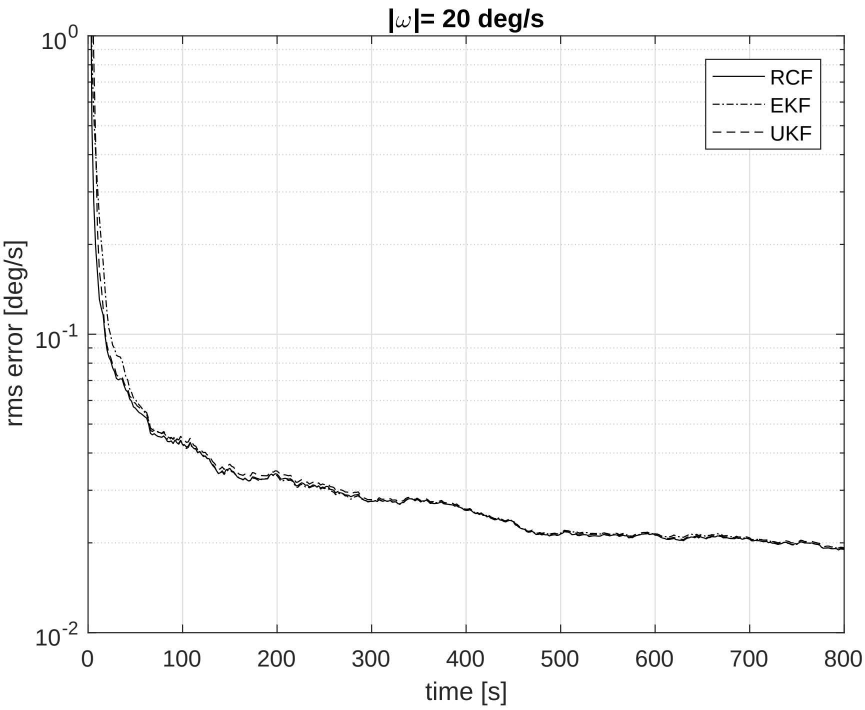

Figure 1 Estimated angular velocity rms error for initial S/C angular speed of 20 deg/s.

VI. Conclusions

In this Note, a nonlinear filtering scheme based on a set of recursive equations was analyzed. The algorithm, named retrodictor-corrector filter (RCF), is similar to existing variants of the Kalman filter, but it is derived under different assumptions, aimed at mitigating the approximation intrinsic to the nonlinear predictor step common to many recursive filters.

The estimation is expressed as a weighted least squares problem, where the cost function reflects the balance between a minimization of the measurement and a-priori residuals. However, rather than fitting the current estimate to a previous one being projected-ahead, as done in the standard predictor-corrector scheme, here we seek to retrofit the current estimate to the previous one. This way, the concept of successive linearization is applied not only to the measurement function, but

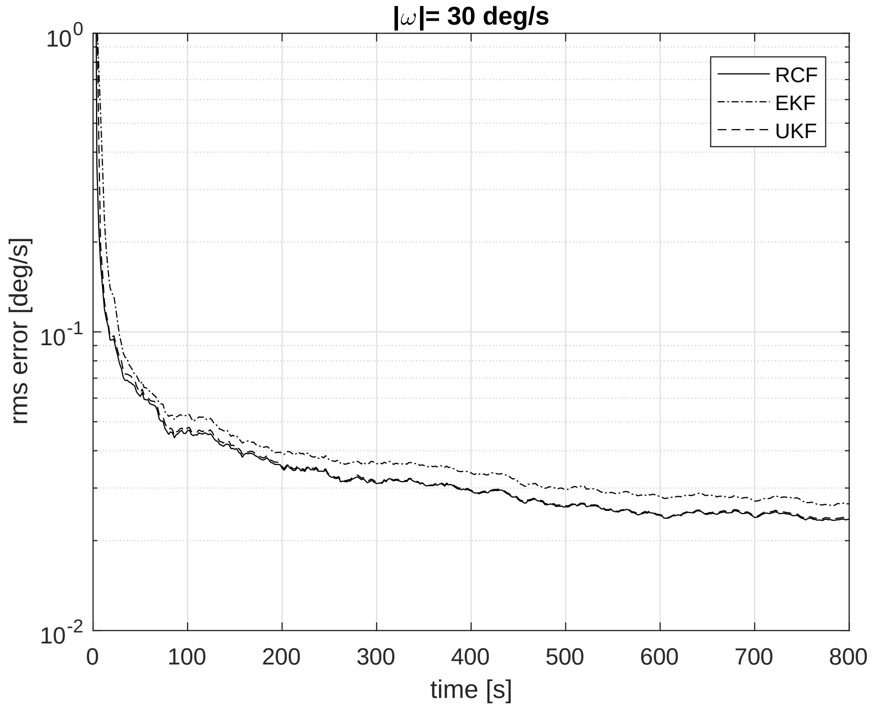

Figure 2 Estimated angular velocity rms error for initial S/C angular speed of 30 deg/s.

also to the dynamic equation. As a consequence, the proposed algorithm effectively handles systems having strongly nonlinear dynamics, without resorting to a smoothing stage. A solution scheme was presented, based on the iterative differential correction method, which was shown to lead at the first iteration to the extended Kalman filter (EKF).

A numerical example was discussed involving the angular rate estimation of a tumbling satellite from magnetometer readings. Results obtained with the proposed algorithm were compared to the ones from an unscented Kalman Filter (UKF) and the EKF: the retrodictor-corrector filter proved to work very well under severe nonlinearities, with performance analogous to the UKF ones.

We do not pretend to break new ground here with this work, since the filter collects and put under a different perspective some existing concepts, nor we argue that the improvement gain with respect to an EKF is necessary in the majority of applications encountered in practice. Nevertheless, the mathematical formulation and the physical interpretation proposed herein prove to be effective in real-time estimation problems. The proposed solution algorithm requires very few modifications with respect to a standard EKF implementation and, in addition, the iteration procedure can easily

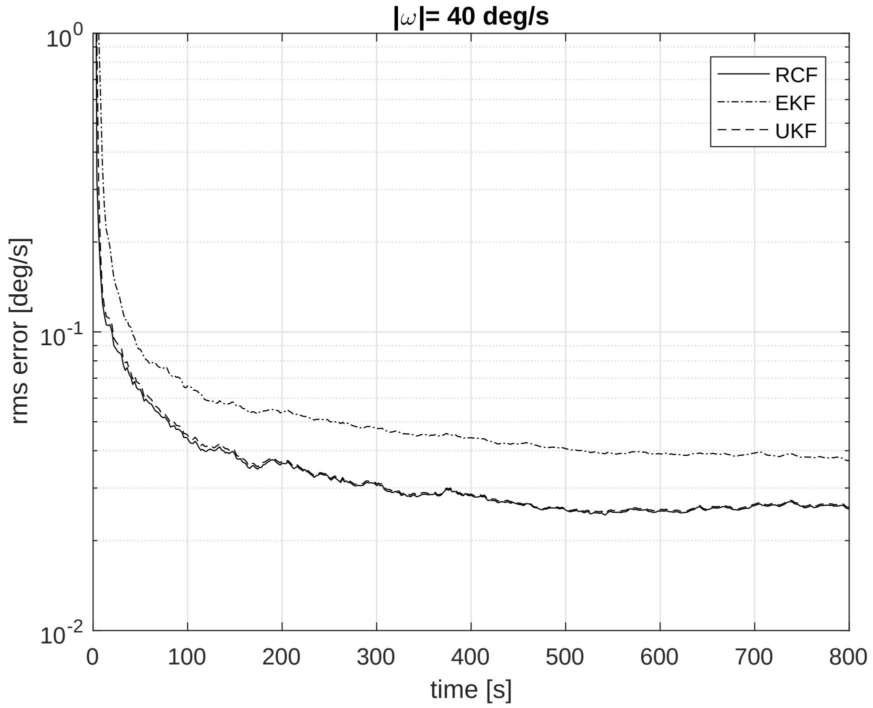

Figure 3 Estimated angular velocity rms error for initial S/C angular speed of 40 deg/s.

self recognize whether or not the problem at hand requires further iterations beyond the EKF one, by simply checking the behaviour of the cost function between successive iterations.

Appendix. Equivalence of the predicted and retrodicted a-priori residuals for linear dynamics

If the system dynamic can be described through:

$$\boldsymbol{x}_k = \boldsymbol{\Phi}_{k,k-1}\boldsymbol{x}_{k-1} + \boldsymbol{\Gamma}_{k-1}\boldsymbol{w}_{k-1} \tag{29}$$

Then it follows:

$$\boldsymbol{x}_{k-1} = \boldsymbol{\Phi}_{k-1,k}\boldsymbol{x}_k - \boldsymbol{\Phi}_{k-1,k}\boldsymbol{\Gamma}_{k-1}\boldsymbol{w}_{k-1} \tag{30}$$

Therefore, the backward process noise reads:

$$\boldsymbol{u}_{k-1,k} = \boldsymbol{\Phi}_{k-1,k}\boldsymbol{\Gamma}_{k-1}\boldsymbol{w}_{k-1} \tag{31}$$

The effective covariance to be used to weight the retrodicted a-priori follows as:

$$\boldsymbol{P}_{k-1,k} = \boldsymbol{P}_{k-1} + \boldsymbol{\Phi}_{k-1,k}\boldsymbol{\Gamma}_{k-1}\boldsymbol{Q}_{k-1}\boldsymbol{\Gamma}_{k-1}^T\boldsymbol{\Phi}_{k-1,k}^T \tag{32}$$

which is a backward propagation of the familiar a-priori covariance in the prediction framework:

$$\boldsymbol{P}_{k|k-1} = \boldsymbol{\Phi}_{k,k-1}\boldsymbol{P}_{k-1}\boldsymbol{\Phi}_{k,k-1}^T + \boldsymbol{\Gamma}_{k-1}\boldsymbol{Q}_{k-1}\boldsymbol{\Gamma}_{k-1}^T = \boldsymbol{\Phi}_{k,k-1}\boldsymbol{P}_{k-1,k}\boldsymbol{\Phi}_{k,k-1}^T \tag{33}$$

Now the equivalence of the first terms in the right hand sides of Eqs. (18) and (19) follows as:

$$\begin{aligned} \|\tilde{\boldsymbol{x}}_{k-1} - \boldsymbol{\Phi}_{k-1,k}\tilde{\boldsymbol{x}}_k\|^2_{\boldsymbol{P}^{-1}_{k-1,k}} &= \\ \|\tilde{\boldsymbol{x}}_{k-1} - \boldsymbol{\Phi}_{k-1,k}\tilde{\boldsymbol{x}}_k\|^2_{\boldsymbol{\Phi}^T_{k,k-1}\boldsymbol{\Phi}^{-T}_{k,k-1}\boldsymbol{P}^{-1}_{k-1,k}\boldsymbol{\Phi}^{-1}_{k,k-1}\boldsymbol{\Phi}_{k,k-1}} &= \\ \|\tilde{\boldsymbol{x}}_{k-1} - \boldsymbol{\Phi}_{k-1,k}\tilde{\boldsymbol{x}}_k\|^2_{\boldsymbol{\Phi}^T_{k,k-1}\left(\boldsymbol{\Phi}_{k,k-1}\boldsymbol{P}_{k-1,k}\boldsymbol{\Phi}^T_{k,k-1}\right)^{-1}\boldsymbol{\Phi}_{k,k-1}} &= \\ \|\tilde{\boldsymbol{x}}_{k-1} - \boldsymbol{\Phi}_{k-1,k}\tilde{\boldsymbol{x}}_k\|^2_{\boldsymbol{\Phi}^T_{k,k-1}\boldsymbol{P}^{-1}_{k|k-1}\boldsymbol{\Phi}_{k,k-1}} = \|\boldsymbol{\Phi}_{k,k-1}\tilde{\boldsymbol{x}}_{k-1} - \tilde{\boldsymbol{x}}_k\|^2_{\boldsymbol{P}^{-1}_{k|k-1}} & \end{aligned} \tag{34}$$

References

[1] Gelb, A., *Applied Optimal Estimation*, MIT Press, 1974, Sec. 6.1, pp. 182–192.

[2] Wishner, R.P. Tabaczynski, J.A. and Athans, M. “A comparison of three non-linear filters” *Automatica*, Vol. 5, No. 4, 1969, pp. 487–496. doi:10.1016/0005-1098(69)90110-1

[3] Psiaki, M. “Backward-Smoothing Extended Kalman Filter”, *Journal of Guidance, Control, and Dynamics*, Vol. 28, No. 5, 2005, pp. 885–894. doi: 2514/1.12108

[4] Van Der Merwe, R. and Wan, E. “Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models”, *Proceedings of the Workshop on Advances in Machine Learning*, 2003.

[5] Julier, S.J. and Uhlmann, J.K. “Unscented filtering and nonlinear estimation," in *Proceedings of the IEEE*, Vol. 92, No. 3, Mar 2004, pp. 401-422. doi: 10.1109/JPROC.2003.823141.

[6] Morrison, N., *Introduction to Sequential Smoothing and Prediction*, McGraw-Hill, 1st edition, 1969.

[7] Biermann G. J., Factorization Methods for Discrete Sequential Estimation, Dover, 2006, Ch. 2, pp. 13–19.

[8] Kelley, C.T., *Iterative Methods for Optimization*, Society for Industrial and Applied Mathematics, 1999, Sec. 2.4, pp. 22–28 doi: 10.1137/1.9781611970920

[9] Hodgart, M.S., Tortora, P., and Morrison, N., “A Gaussian Deterministic Approach to Smoothing and Prediction for Application to Spacecraft Attitude and Orbit Determination,” Proceedings of the National Aerospace & Electronics Conference (NAECON 2000), Dayton, Ohio, 10–12 Oct. 2000, pp. 687–695. doi: 10.1109/NAECON.2000.894980

[10] de Angelis, E.L., Ferrarese, G., Giulietti, F., Modenini, D., and Tortora, P., “Terminal height estimation using a Fading Gaussian Deterministic filter”, *Aerospace Science and Technology*, Vol. 55, August 2016, pp. 366-376. doi: 10.1016/j.ast.2016.06.013.

[11] Tortora, P., Oshman, Y., and Santoni, F., "Spacecraft Angular Rate Estimation from Magnetometer Data Only Using an Analytic Predictor", *Journal of Guidance, Control, and Dynamics*, Vol. 27, No. 3, May–June 2004, pp. 365–373. doi: 10.2514/1.10332

[12] de Angelis, E.L., Ferrarese, G., Giulietti, F., Modenini, D., and Tortora, P., "Gaussian Deterministic Recursive Estimator with Online Tuning Capabilities", *Journal of Guidance, Control, and Dynamics*, Vol. 38, No. 9., pp. 1827–1833. doi: 10.2514/1.G001024.

[13] Steyn, W. H., Hashida, Y. and Lappas, V. "An Attitude Control System and Commissioning Results of the SNAP-1 Nanosatellite", *Proceedings of the 14th Annual AIAA/USU Conference on Small Satellites*, Utah State University, Logan Utah USA, 10–12 Oct. 2000.

[14] Rossi, F., Ivanov, A., Muriel, R. and Gass, V. "The SwissCube technologies results after four years of flight", *Proceedings of the 63th International Astronautical Congress*, Beijing, China, 23–27 Sep. 2013.

[15] These are not to be confused with *smoothed* estimates, usually denoted in the estimation literature as $\tilde{\boldsymbol{x}}_{k-1|k}$ and $\boldsymbol{P}_{k-1|k}$.

[16] In other words: $\tilde{\boldsymbol{x}}_{k-1,k} = \tilde{\boldsymbol{x}}_k + \int_{t_k}^{t_{k-1}} \boldsymbol{f}(\tilde{\boldsymbol{x}}, \tau) d\tau$

[17] Using a standard nomenclature: $\alpha = 10^{-3}$, $\kappa = 0$, $\beta = 2$.