ESTIMATION OF THE CONDITIONAL STATE AND COVARIANCE WITH TAYLOR POLYNOMIALS

Simone Servadio,[*] Renato Zanetti[†]

A novel estimator is presented that expands the typical state and covariance update laws of Kalman filters to polynomial updates in the measurement. The filter employs Taylor series approximations of the nonlinear dynamic and measurement functions. All polynomials (functions approximation, state update, and covariance update) can be made to arbitrary order to trade between filter's accuracy/consistency and computational time. The performance of the algorithm is tested in numerical simulations.

INTRODUCTION

Estimation is the process of inferring the value of a quantity of interest from indirect, inaccurate and noisy observations. When the quantity of interest is the (current) state of a dynamic system, the problem is referred to as "filtering": the best estimate is obtained by "filtering out" the noise from noisy measurements. The estimate is the output given by an optimal estimator, which is a computational algorithm that processes measurements while maximizing a certain performance index. The optimal estimator makes the best use of the data, the knowledge of the system, and the disturbances.

State estimation has a great variety of applications. Among the most important are tracking, surveillance, trajectory determination, navigation, control theory (such as guidance, attitude control, sensor pointing,...), signal processing, image processing, communications, mapping, and etc. Most of these problems, if not all, share nonlinear dynamics systems, usually observed via nonlinear measurements. While optimal estimators for linear systems are known analytically in closed form, analogous results do not exist for estimation of nonlinear systems. Thus, a closed form solution of the nonlinear estimation problem is typically unavailable, and assumptions are typically made to make the problem tractable. As a consequence the state estimate of nonlinear systems is suboptimal.

Different algorithms that can deal with nonlinearities have been developed to achieve an accurate estimate. They use various approximations of the probability density functions (through samplings, moments, polynomials, kernels, ...) and of their propagation in time (such as state transition tensors, the unscented transformation, polynomials, ...). Consequently, most current estimators provide an estimate that is a function of the measurement outcome (linearly or nonlinearly), where the estimated covariance is the mean among all of its possible realizations. The covariance is not directly influenced by the measurement itself, but it depends on the selected approximation.

The filtering problem is an important research area that has attracted considerable interest, especially for the challenging case of nonlinear dynamic systems, where the estimation of the state of the system from noisy data cannot usually be solved analytically. For the well-known linear and Gaussian case, the posterior distribution remains Gaussian and the Kalman Filter [1, 2] provides the mechanization to calculate its mean and covariance. However, most problems are nonlinear in the dynamics and in the measurement equations, corresponding to a non-Gaussian posterior probability density function (PDF). Most applications in aerospace engineering deal with high nonlinearities: orbit determination [3], spacecraft navigation, target tracking, etc., require a reliable nonlinear filtering technique able to approximate the non-Gaussian posterior distribution.

[*]Ph.D. Candidate, Aerospace Engineering and Engineering Mechanics, The University of Texas at Austin
[†]Assistant Professor, Aerospace Engineering and Engineering Mechanics, The University of Texas at Austin

Many techniques have been developed to deal with the nonlinear estimation problem. The first solution is based on the linearization of the dynamics and measurement equations around the most current estimate. The Extended Kalman Filter (EKF) [4] algorithm applies the Kalman filter mechanization to the linearized system, where PDFs are approximated as Gaussians (and they remain Gaussian under a liner transformation). However, in problems with high nonlinearities, such as orbit determination, the linearization assumption may fail to give a valid estimate [5]. The linearization assumption fails for high initial uncertainties levels or long propagation periods.

Another well-know technique to account for the system nonlinearities is the unscented transformation. The Unscented Kalman filter (UKF) [6, 7] is able to better handle the effects of nonlinearities in the dynamics and in the measurements and, typically, it achieves higher accuracy and robustness levels when compared to the EKF. The UKF applies the unscented transformation to achieve a more accurate Gaussian approximation of the predicted PDFs. The UKF is a linear estimator, i.e. the estimate is a linear function of the current measurement.

The first order approximation of the EKF can e extended to higher order Taylor series [4, 8]. Generally, the higher the order of the Taylor series, the better the performance of the filter. The Gaussian Second Order Filter (GSOF) [9] truncates the Taylor series at second order to achieve a more accurate estimate. Truncating the Taylor series to order c requires knowledge of the estimation error's central moments up to order $2c$ in order to calculate the Kalman gain, e.g. the EKF truncates at first order and it requires knowledge of the covariance matrices. Consequentially, the GSOF requires knowledge of the third and forth central moments of the state distribution. The GSOF approximates the PDF as Gaussian and the high order moments are calculated directly from the covariance matrix. The GSOF performs a linear update based on a second order approximation of the posterior estimation error. Gaussian filters exist up to any arbitrary truncation order of the Taylor series expansion using Differential Algebra (DA) techniques [10].

Park and Scheeres [11] use state transition tensors (STT) to propagate mean and higher order central moments through arbitrary nonlinear dynamics. Their filtering technique better incorporates the nonlinear dynamics during the prediction step of the algorithm, and keeps a linear update structure. Their work has been subsequently expanded to create higher order Kalman filters able to handle process noise and measurement updates [12]. Their measurement update only involves the mean and the covariance of the state, neglecting the contribution on the higher order central moments. Majji, Turner, and Junkins [13], on the other hand, introduce a tensorial mechanization that expands the work by Park and Scheeres to include measurement updates in all the higher order central moments. Their algorithm has been replicated in the DA framework by Valli *et al.* [14], achieving the same results with an efficient mechanization.

All of the filters mentioned above are linear estimators, i.e. the estimate is a linear function of the measurements. The conditional mean, which is the optimal Minimum Mean Square Error (MMSE) solution, is typically some unknown nonlinear function of the measurement whose computation is usually not feasible. A linear estimator, even when accounting for the nonlinearities of the measurement function, is typically outperformed by nonlinear estimators such as the Gaussian Sum Filter (GSF) [15, 16] or the (Bootstrap) Particle Filter (BPF) [17].

Gaussian Sum Filters (GSF) approximate PDFs with multiple Gaussian kernels [18] and apply a linear filter's equations to each model (typically EKF or UKF). The GSF divides the state uncertainties into multiple smaller subdomains where the linear approximation of the functions is more accurate. The Gaussian Multiple Model (GMM) final estimate is a weighted sum of the estimates from each model, based on the measurement likelihood function. The time propagation of the GSF can fail to represent the propagated PDF correctly in the presence of sporadic measurements when component's covariances grow too large. Terejanu *et al.* [19] improve the GFS' propagation accuracy by adapting the weights of each component during the propagation phase. DeMars *et al.* [20] address the nonlinearities of the dynamics by refining the number of mixand components.

Raihan and Chakravorty [21] developed a particle Gaussian mixture filter robust to extremely high initial level of uncertainties. Valli et al. [22] use differential algebra (DA) techniques to develop a Monte Carlo Kalman filter that substitutes samples propagation with polynomial evaluations. Servadio and Zanetti pre-

sented a filter [23] that exploit the DA propagation from Valli [22] and mixes it with the GMM update from Chakravorty [21].

Other techniques to develop a nonlinear estimator have been presented. De Santis *et al.* [24] propose an augmented state to obtain a nonlinear update but preserving the linear update structure. Their work, focused on linear and non-gaussian systems, offers an approximation of the the optimal non-linear update as either quadratic [24] or polynomial [25].

Servadio and Zanetti [26] also implemented a quadratic update (extendable to polynomial update of any order) based on Taylor series expansions, implemented in the DA framework. The polynomial representation of the state of the system allows the evaluation of central moments up to a selected order: the polynomial high-order coefficients of the estimator are evaluated to minimize the mean square estimation error. The proposed filter is robust against non-Gaussian noises and non-Gaussian initial conditions, but the computational demand of calculating higher order central moments can grow quickly with the truncation order of the Taylor series, the size of the state vector, and the order of the polynomial update. Servadio and Zanetti also developed a new technique [27] where non-Gaussian distributions are approximated as a polynomial transformation of Gaussian random variables. In doing so, all high order central moments are easily and efficiently calculated in close form. Consequently, in [27], polynomial updates can be performed much more efficiently than in [26]. The algorithm can be interpreted as an expansion of Gaussian filters, where the shape of the distributions is approximated through a polynomial function of Gaussian random vectors.

The paper is structured in the following way. At first, the aim of the manuscript is highlighted, focusing on the innovative contributions and introducing differential algebra (DA) techniques. The following chapter describes the polynomial estimator and its mathematical derivation. The main part of the paper offers a new filtering technique developed in the differential algebra framework. Later, the new algorithm is applied to two numerical examples: a scalar application, in order to graphically visualize how the new algorithm works, and the Lorenz96 dynamics Lastly, conclusions are drawn.

BACKGROUND

Filters fall under either of two categories: linear and nonlinear. Both approaches use Bayes' rule for the measurement update, but while the latter typically parameterizes the complete conditional PDF, the former usually only carries mean and covariance to fully express the PDF. The linear update rule for mean and covariance are given by

$$\hat{\mathbf{x}}^+ = \hat{\mathbf{x}}^- + \mathbf{K}(\mathbf{y} - \hat{\mathbf{y}}^-) \tag{1}$$

$$\mathbf{\Sigma}_{\mathbf{xx}}^+ = \mathbf{\Sigma}_{\mathbf{xx}}^- - \mathbf{K}\mathbf{\Sigma}_{\mathbf{yy}}\mathbf{K}^{\mathbf{T}} \tag{2}$$

where $\mathbf{K}$ is the Kalman gain, $\hat{\mathbf{y}}^-$ is the predicted measurement mean, and $\hat{\mathbf{x}}^-$ is the prior mean. The equations above are only optimal in a minimum mean square error (MMSE) when the prior distribution and the measurement are jointly Gaussian (which implies a linear relation between the two). In general the MMSE estimate is the conditional mean, an unknown and typtically nonlinear functino of the measurement outcome $\mathbf{y}$; Eq. (1) is the statistical linearization regression of the conditional mean [28], that is to say: equation Eq. (1) is the best linear fit approximation of the conditional mean. Eq. (2) on the other hand, is the best constant approximation of the conditional covariance, also in the statistical regression sense.

For nonlinear measurement functions, $\hat{\mathbf{y}}^-$, $\mathbf{\Sigma}_{\mathbf{xx}}^-$, e $\mathbf{\Sigma}_{\mathbf{yy}}$ are typically not obtainable in closed form, and different Kalman filter algorithms make different approximations to calculate them: Jacobians (EKF)[4], unscented transformation (UKF)[6], polynomials (High Order KF, [14]), Gaussian quadrature (QKF)[29], spherical cubature (CKF)[30], Monte Carlo (MCKF), ensemble (EnKF)[22], central differences (CDKF)[31], finite differences (DDKF) [32], etc..

Regarding the second case, the estimator can achieve nonlinearity in different ways: through multiple models, such as GSF [18] with GMM, sequential Monte Carlo, such as Bootstrap PF (BPF) [17], Marginalized PF (MPF) [33], Auxiliary PF (APF), Unscented PF (UPF) [34], Gaussian PF (GPF), Monte Carlo Filter PF (MCFPF) [35], or polynomial update [26]. However, for multiple models and particle filters, the estimated

covariance is a function of the measurements, since it is evaluated as a weighed mean based on the measurement likelihood. The drawback of these approaches is the high computational effort requested by the filter from the machine, which has to work with an elevated number of particles or models to achieve a highly precise prediction of the conditional covariance.

Experience has shown that the order of the statistical regression approximation of the covariance needs to be lower than that of the mean for good numerical performance of the algorithm. A zeroth order covariance approximation, therefore, has endured as a companion of a linear mean update rule. Our prior work [27] presents a higher order polynomial state update, i.e. a higher order polynomial approximation of the conditional mean. With this paper, we develop a novel higher order polynomial covariance update to better approximate the conditional covariance than the standard zeroth order approach. Our polynomial estimator is implemented in the Differential Algebra framework, and the next two subsections briefly recap the two.

Differential Algebra

The theory of differential algebra (DA) has been developed by Martin Berz in the late 80's [36]. The DA framework [37, 8] benefits of an algebra of Taylor polynomials. Therefore, each function is represented through a matrix of coefficients and exponents instead than the classical representation with an array of floating points (FP) numbers. The Differential Algebra Core Engine (DACE2.0) [38, 39] software has an hard-coded library of the Taylor expansion series of elementary functions. As a consequence, derivatives are not computed numerically (e.g finite differences), but evaluated directly from the Taylor polynomials. Therefore, DA offers a new way of working in a computer environment where the new algebra of polynomials is endowed of composition of function, function inversions, explicit system solving, etc., likewise the classic FP arithmetic.

Differential algebra has been proven to reduce computational costs in solving ordinary differential equations (ODE) [40, 10]. Once the maximum truncation order of the polynomial has been selected, DA creates the Taylor polynomial expansion of the flow of ODES as a function of the provided initial conditions. This approach can replace thousands of integrations with the computationally faster evaluation of the Taylor expansion [41, 23]. As a results, the computational burden requested from the machine reduces considerably [22]. In the filtering problem, DA techniques have been used for the development of an efficient mapping of uncertainties [14], for the evaluation of high-order moments [42], and for the approximation of the shape of PDFs [43]. Wittig *et al.* [44] developed a domain splitting technique that improves the state propagation when initial uncertainties are large by creating multiple polynomials.

The main concept of DA is that each function $f(\mathbf{x})$ can be expressed as a polynomials $p(\delta\mathbf{x})$; where the new variable $\delta\mathbf{x}$ is the deviation from the expansion center $\hat{\mathbf{x}}$. The polynomial $p(\delta\mathbf{x})$ is the Taylor expansion series of $f(\mathbf{x})$, centered in $\hat{\mathbf{x}}$, and truncated up to an user-selected order c.

For a detailed description of DA, its techniques, and how the DACE2.0 works in a computer environment, the reader is advised to refer the references.

The Polynomial Estimator

Our previous work [27] presented a filter that uses polynomial transformations of Gaussian random variables in an augmented state to represent non-Gaussian distributions via their high order moments. The filtering technique presented in this paper uses the same representation of variables, but extend its contribution to further improve the estimation of the system's uncertainties.

Let $\mathbf{x}$ be the state of the system which is desired to estimate, and let $\mathbf{y}$ be another random vector, sampleable, related to $\mathbf{x}$. State estimators consist in a family of functions $\mathbf{g}(\mathbf{y})$ that infer the unknown value of $\mathbf{x}$ based on the know outcome $\mathbf{y}$. Polynomial estimators are a subset of this family. Using to the Kronecker algebra, the polynomial estimator, up to an arbitrary order, can be written as

$$\mathbf{g}(\mathbf{y}) = \mathbf{a} + \mathbf{K}_1\mathbf{y} + \mathbf{K}_2\mathbf{y}^{[2]} + \mathbf{K}_3\mathbf{y}^{[3]} + \mathbf{K}_4\mathbf{y}^{[4]} + \dots \tag{3}$$

4

where $\mathbf{a}$ is a constant, each $\mathbf{K}_i$ is a constant matrix of appropriate dimensions, and each $\mathbf{y}^{[i]}$ is calculated using the Kronecker product ($\otimes$)

$$\mathbf{y}^{[2]} = \mathbf{y} \otimes \mathbf{y} \tag{4}$$

$$\mathbf{y}^{[i]} = \mathbf{y} \otimes \mathbf{y} \otimes \mathbf{y} \otimes \dots \tag{5}$$

In order to avoid redundancy, each repeated component is eliminated; thus, in $\mathbf{y}^{[2]}$, only one term between $y_1 y_2$ and $y_2 y_1$ is kept. Consistently with previous works, it is convenient to derive the estimator's constants by working with deviation vectors. It is easier to define central moments of the deviation vectors of the variable from their mean than of the vectors themselves. Therefore, following the notation applied in [27], deviation vectors are defined as

$$d\mathbf{x} = \mathbf{x} - \mathbb{E}\{\mathbf{x}\} \tag{6}$$

$$d\mathbf{y} = \mathbf{y} - \mathbb{E}\{\mathbf{y}\} \tag{7}$$

$$d\mathbf{y}^{\{2\}} = \mathbf{y} \otimes \mathbf{y} - \mathbb{E}\{\mathbf{y} \otimes \mathbf{y}\} = \mathbf{y}^{[2]} - \mathbb{E}\left\{\mathbf{y}^{[2]}\right\} \tag{8}$$

$$d\mathbf{y}^{\{i\}} = \mathbf{y}^{[i]} - \mathbb{E}\left\{\mathbf{y}^{[i]}\right\} \tag{9}$$

where Equation (6) is the state deviation and Equation (7) is the measurement residual. It is worth noticing that deviations have zero mean by construction. The family of polynomial estimators defined by Equation (3) can be redefined by adding and subtracting constants, in order to obtain a new, but theoretically equivalent, family

$$\mathbf{g}(\mathbf{y}) = \mathbf{a} + \mathbb{E}\{\mathbf{x}\} + \mathbf{K}_1(\mathbf{y} - \mathbb{E}\{\mathbf{y}\}) + \mathbf{K}_2\left(\mathbf{y}^{[2]} - \mathbb{E}\left\{\mathbf{y}^{[2]}\right\}\right) + \mathbf{K}_3\left(\mathbf{y}^{[3]} - \mathbb{E}\left\{\mathbf{y}^{[3]}\right\}\right) + \dots \tag{10}$$

$$= \mathbf{a} + \mathbb{E}\{\mathbf{x}\} + \mathbf{K}_1 d\mathbf{y} + \mathbf{K}_2\, d\mathbf{y}^{\{2\}} + \mathbf{K}_3\, d\mathbf{y}^{\{3\}} + \dots \tag{11}$$

The measurement residual and its powers can be stacked, defining the augmented deviation vector

$$d\mathcal{Y} = \begin{bmatrix} d\mathbf{y}^{\mathrm{T}} & d\mathbf{y}^{\{2\}\mathrm{T}} & d\mathbf{y}^{\{3\}\mathrm{T}} & \dots \end{bmatrix}^{\mathrm{T}} \tag{12}$$

The same reasoning applies for the constants

$$\mathcal{K} = \begin{bmatrix} \mathbf{K}_1 & \mathbf{K}_2 & \mathbf{K}_3 & \dots \end{bmatrix} \tag{13}$$

such that the generic polynomial estimator illustrated by Equation (11) can be rewritten as

$$\mathbf{g}(\mathbf{y}) = \mathbf{a} + \mathbb{E}\{\mathbf{x}\} + \mathcal{K}d\mathcal{Y} \tag{14}$$

The optimal estimator, in a Minimum Mean Square Error (MMSE) sense, satisfies the orthogonality principle (full proof in [26, 27]), which gives the optimal coefficients. The optimal polynomial update estimator becomes

$$\hat{\mathbf{x}} = \mathbb{E}\{\mathbf{x}\} + \mathbf{P}_{\mathbf{x}\mathcal{Y}}\mathbf{P}_{\mathcal{Y}\mathcal{Y}}^{-1}d\mathcal{Y} \tag{15}$$

where $\hat{\mathbf{x}}$ is the estimate. Matrices $\mathbf{P}_{\mathbf{x}\mathcal{Y}}$ and $\mathbf{P}_{\mathcal{Y}\mathcal{Y}}$ are, respectively, the augmented state-measurement cross-covariance matrix and the augmented measurement covariance matrix, defined blockwise as

$$\mathbf{P}_{\mathbf{x}\mathcal{Y}} = \begin{bmatrix} \mathbf{P}_{\mathbf{x}\mathbf{y}} & \mathbf{P}_{\mathbf{x}\mathbf{y}^{[2]}} & \mathbf{P}_{\mathbf{x}\mathbf{y}^{[3]}} & \dots \end{bmatrix} \tag{16}$$

$$\mathbf{P}_{\mathcal{Y}\mathcal{Y}} = \begin{bmatrix} \mathbf{P}_{\mathbf{y}\mathbf{y}} & \mathbf{P}_{\mathbf{y}\mathbf{y}^{[2]}} & \mathbf{P}_{\mathbf{y}\mathbf{y}^{[3]}} & \dots \\ \mathbf{P}_{\mathbf{y}^{[2]}\mathbf{y}} & \mathbf{P}_{\mathbf{y}^{[2]}\mathbf{y}^{[2]}} & \mathbf{P}_{\mathbf{y}^{[2]}\mathbf{y}^{[2]}} & \dots \\ \mathbf{P}_{\mathbf{y}^{[3]}\mathbf{y}} & \mathbf{P}_{\mathbf{y}^{[3]}\mathbf{y}^{[2]}} & \mathbf{P}_{\mathbf{y}^{[3]}\mathbf{y}^{[3]}} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{17}$$

where the following notation is used:

$$\mathbf{P}_{\mathbf{x}\mathbf{y}} = \mathbb{E}\left\{(\mathbf{x} - \mathbb{E}\{\mathbf{x}\})(\mathbf{y} - \mathbb{E}\{\mathbf{y}\})^{\mathrm{T}}\right\} \tag{18}$$

Since deviations have zero mean by construction, the identities $\mathbf{P}_{\mathbf{y}^{[i]}\mathbf{y}^{[j]}} = \mathbf{P}_{d\mathbf{y}^{\{i\}}d\mathbf{y}^{\{j\}}}$ and $\mathbf{P}_{\mathbf{x}\mathbf{y}^{[j]}} = \mathbf{P}_{d\mathbf{x}d\mathbf{y}^{\{j\}}}$ are valid $\forall i, j \in \mathbb{N}_0$.

THE STATE AND COVARIANCE ESTIMATION FILTER

The state of the art has pointed out the lack in literature of a filter that estimates the state uncertainty level with the same importance reserved for the state itself. A new filtering technique, based on a double polynomial estimator, have been developed in the DA framework. The double nature of the filter refers to the sequential estimation of the state and the covariance, where, at each time step, the same measurement outcome is used twice to achieve matching between the conditioned state mean and its relative uncertainty spread.

Consider the generic dynamic system described by the following equations of motion and measurement equations:

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k) + \mathbf{v}_k \tag{19}$$

$$\mathbf{y}_{k+1} = \mathbf{h}(\mathbf{x}_{k+1}) + \mathbf{w}_{k+1} \tag{20}$$

where $\mathbf{f}(\cdot)$ is the dynamics function, $\mathbf{x}_k$ is the n-dimensional state of the system at time-step k, $\mathbf{y}_{k+1}$ is the m-dimensional measurement vector at time-step $k+1$, and $\mathbf{h}(\cdot)$ is the measurement function. The noises are assumed to be zero mean Gaussians and uncorrelated, such that their distribution is fully described by the first two moments. For all discrete time indexes i and j

$$\mathbb{E}\left\{\mathbf{v}_i \mathbf{w}_j^T\right\} = 0 \tag{21}$$

$$\mathbb{E}\left\{\mathbf{v}_i \mathbf{v}_j^T\right\} = \mathbf{Q}_i \delta_{ij} \tag{22}$$

$$\mathbb{E}\left\{\mathbf{w}_i \mathbf{w}_j^T\right\} = \mathbf{R}_i \delta_{ij} \tag{23}$$

where $\mathbf{Q}_i$ is the process noise autocovariance function while $\mathbf{R}_i$ is for the measurement noise. The initial condition of the state of the system is assumed to be Gaussian as well $\mathbf{x}_0 \sim \mathcal{N}(\hat{\mathbf{x}}_0, \mathbf{P}_0)$; however, for all other time steps $k > 0$, the state distribution will be non-Gaussian due the high nonlinearities in the equations of motion.

The State And Covariance Estimation Filter, SACE-c-η-μ, is composed of three different parts: the prediction, the state update, and the covariance update. The three integers c, η and μ in SACE-c-η-μ refers to the tuning parameters of the filter: they are, respectively, the order of the Taylor polynomial expansion (c), the order of the state polynomial update (η), and the order of the covariance polynomial update (μ).

Prediction

At the beginning of each time step, the state distribution is assumed to be Gaussian $\mathbf{x}_k \sim \mathcal{N}(\hat{\mathbf{x}}_{\mathbf{k}}, \mathbf{P}_k)$. The state can therefore be initialized in the DA framework as a first order polynomial

$$\mathbf{x}_k = \mathbf{x}_k(\delta\mathbf{x}_k) = \hat{\mathbf{x}}_{\mathbf{k}} + \mathbf{S}_k \delta\mathbf{x}_k \tag{24}$$

where $\mathbf{S}_k \mathbf{S}_k^T = \mathbf{P}_k$ and the DA variable $\delta\mathbf{x}_k = \mathbf{x}_k - \hat{\mathbf{x}}_{\mathbf{k}}$ expresses the deviation from the expansion center and it is interpreted as a Gaussian with zero mean and identity covariance matrix. Therefore, matrix $\mathbf{S}_k$ (here calculated through Cholesky Decomposition), scales the coefficients of the state polynomial such that the moments of $\mathbf{x}_k$ can be calculated directly from the moments of $\mathcal{N}(\mathbf{0}, \mathbf{I})$.

The propagation function is applied directly on the state polynomial, such that the predicted state vector is

$$\mathbf{x}_{k+1}^- = \mathbf{x}_{k+1}^-(\delta\mathbf{x}_k) = \mathbf{f}\left(\mathbf{x}_k(\delta\mathbf{x}_k)\right) \tag{25}$$

where $\mathbf{x}_{k+1}^-$ indicates the Taylor expansion series of the dynamics centered in $\hat{\mathbf{x}}_k$ truncated at the user-defined integer order c. Equation (25) is carried out in the DA framework, where all the coefficients of the polynomial expansions are evaluated through partial derivatives [8, 26]. Therefore, each component of $\mathbf{x}_{k+1}^-$ is a polynomial map (centered in $\hat{\mathbf{x}}_k$) that maps deviations $(\delta\mathbf{x}_k)$ from time step k to time step $k+1$ and describes how the state PDF evolves in time [42]. The predicted polynomials are lacking the influence of the process noise. Process noise can be mapped in the DA framework with the same representation reserved for

the state of the system. Thus, a new DA variable $\delta\mathbf{v}_k$, interpreted again as a standard normal random vector, is introduced

$$\mathbf{x}_{k+1}^-(\delta\mathbf{x}_k, \delta\mathbf{v}_k) := \mathbf{x}_{k+1}^-(\delta\mathbf{x}_k) + \mathbf{T}_k\delta\mathbf{v}_k \tag{26}$$

where $\mathbf{v}_k = \mathbf{T}_k\delta\mathbf{v}_k$ and $\mathbf{T}_k\mathbf{T}_k^{\mathrm{T}} = \mathbf{Q}_k$.

Analogously, the predicted measurement are expressed as a Taylor polynomial expansion in the DA framework

$$\mathbf{y}_{k+1} = \mathbf{y}_{k+1}(\delta\mathbf{x}_k, \delta\mathbf{v}_k) = \mathbf{h}\big(\mathbf{x}_{k+1}^-(\delta\mathbf{x}_k, \delta\mathbf{v}_k)\big) \tag{27}$$

where $\mathbf{y}_{k+1}$ is, again, a polynomial centered in $\hat{\mathbf{x}}_k$ with maximum order c. In Equation (27), derivatives are now computed w.r.t. both the state deviation vector $(\delta\mathbf{x}_k)$ and the process noise $(\delta\mathbf{v}_k)$. The influence of the process noise is added to the polynomials likewise in Equation (26): thus, a new DA variable $\delta\mathbf{w}_{k+1}$ is introduced

$$\mathbf{y}_{k+1}(\delta\mathbf{x}_k, \delta\mathbf{v}_k, \delta\mathbf{w}_{k+1}) := \mathbf{y}_{k+1}(\delta\mathbf{x}_k, \delta\mathbf{v}_k) + \mathbf{U}_{k+1}\delta\mathbf{w}_{k+1} \tag{28}$$

where $\mathbf{w}_k = \mathbf{U}_k\delta\mathbf{w}_k$ and $\mathbf{U}_k\mathbf{U}_k^T = \mathbf{R}_k$. One again, $\delta\mathbf{w}_{k+1}$ is interpreted as a standard normal random vector.

All the predicted quantities have been calculated and they are represented as polynomial functions of standard random vectors. The number of variables is $2n + m$: n deviations map the state behavior, n map the process noise, and the remaining m map the measurement noise. The Gaussian nature of the random vectors leads to a fast evaluation of all expectation operations since, for a Gaussian PDF, central moments can be easily computed using the Isserlis' formulation [45].

The State Polynomial Update

The second part of SACE-c-η-μ is the state polynomial update. Thus, after selecting the integer c in the prediction step, the user defines a second integer, η, which selects the order of the polynomial estimator dedicated to the state of the system.

The polynomial update asks for the evaluation of the augmented Kalman gain and for high powers of the measurement polynomials, as expressed in Equation (15). Starting from the latter,

$$\mathbf{y}_{k+1}^{[2]} = \mathbf{y}_{k+1} \otimes \mathbf{y}_{k+1} \tag{29}$$

$$\mathbf{y}_{k+1}^{[i]} = \mathbf{y}_{k+1} \otimes \mathbf{y}_{k+1} \otimes \ldots \tag{30}$$

with $i = 1, \ldots, \eta$ and, once again, the redundant components are eliminated.

The means of the predicted state polynomials are now evaluated. Each polynomial undergoes the expectation operator which, being a linear operator, works directly on the single monomials [10].

$$\hat{\mathbf{x}}^- = \mathbb{E}\left\{\mathbf{x}_{k+1}^-\right\} \tag{31}$$

The deviations have a Gaussian distribution with zero mean and identity covariance matrix, therefore the expected value substitutes the relative Isserlis' moment in for each monomial, according to Table 1. As an

Table 1. Isserlis' moments of Gaussian $\mathcal{N}(0,1)$

exponent	0	1	2	3	4	5	6	7	8	...
coefficient	1	0	1	0	3	0	15	0	105	...

illustrative example: $\mathbb{E}\left\{\alpha\delta x_1^8 \delta x_2^4 \delta x_4^6 \delta v_2^2 \delta w_3^4\right\} = 4725\alpha$. The predicted means of the measurement polynomials are evaluated likewise Equation (31)

$$\hat{\mathbf{y}}_{k+1} = \mathbb{E}\left\{\mathbf{y}_{k+1}\right\} \tag{32}$$

$$\hat{\mathbf{y}}_{k+1}^{[2]} = \mathbb{E}\left\{\mathbf{y}_{k+1}^{[2]}\right\} \tag{33}$$

$$\hat{\mathbf{y}}_{k+1}^{[i]} = \mathbb{E}\left\{\mathbf{y}_{k+1}^{[i]}\right\} \tag{34}$$

where, once again, $i = 1, \ldots, \eta$.

The augmented measurement covariance $\mathbf{P}_{\mathcal{YY},[\eta]}$ is evaluated blockwise according to Equation (17). The matrix is guaranteed to be nonsingular because redundant rows and columns have been eliminated. The matrix is symmetric and each block is evaluated as

$$\mathbf{P}_{\mathbf{y}^{[i]}\mathbf{y}^{[j]}} = \mathbb{E}\left\{(\mathbf{y}_{k+1}^{[i]} - \hat{\mathbf{y}}_{k+1}^{[i]})(\mathbf{y}_{k+1}^{[j]} - \hat{\mathbf{y}}_{k+1}^{[j]})^{\mathrm{T}}\right\} \quad \forall i, j = 1, \ldots, \eta \tag{35}$$

Every time a polynomial multiply himself, the maximum truncation order of the Taylor series doubles. For example, the evaluation of $\mathbf{P}_{\mathbf{y}^{[5]}\mathbf{y}^{[3]}}$ applies the expectation operator to a polynomial with monomials up to order $8c$. The augmented state-measurement cross covariance matrix $\mathbf{P}_{\mathbf{x}\mathcal{Y},[\eta]}$ is evaluated blockwise according to Equation (16): each block is evaluated as

$$\mathbf{P}_{\mathbf{x}\mathbf{y}^{[i]}} = \mathbb{E}\left\{(\mathbf{x}_{k+1}^{-} - \hat{\mathbf{x}}_{k+1}^{-})(\mathbf{y}_{k+1}^{[i]} - \hat{\mathbf{y}}_{k+1}^{[i]})^{\mathrm{T}}\right\} \quad \forall i = 1, \ldots, \eta \tag{36}$$

From these covariances it is now possible to evaluate the augmented Kalman gain

$$\mathcal{K} = \mathbf{P}_{\mathbf{x}\mathcal{Y},[\eta]}\mathbf{P}_{\mathcal{YY},[\eta]}^{-1} \tag{37}$$

The subscript $[\eta]$ specifies that the covariance matrices are created with measurement powers up to order η.

Given $\tilde{\mathbf{y}}_{k+1}$ the numerical outcome of the random vector $\mathbf{y}_{k+1}$, thus the value of measurement coming from the sensors, its powers are evaluated using the Kronecker product

$$\hat{\mathbf{y}}_{k+1}^{[2]} = \tilde{\mathbf{y}}_{k+1} \otimes \tilde{\mathbf{y}}_{k+1} \tag{38}$$

$$\hat{\mathbf{y}}_{k+1}^{[i]} = \tilde{\mathbf{y}}_{k+1} \otimes \tilde{\mathbf{y}}_{k+1} \otimes \ldots \tag{39}$$

with $i = 1, \ldots, \eta$ and, once again, the redundant components are eliminated. The polynomial update exploits the influence of high powers from the measurement outcome. Therefore, the classic measurement residual is developed to create the augmented innovation vector

$$d\tilde{\mathcal{Y}}(\delta\mathbf{x}_k, \delta\mathbf{v}_k, \delta\mathbf{w}_{k+1}) = \begin{bmatrix} \tilde{\mathbf{y}}_{k+1} - \mathbf{y}_{k+1}(\delta\mathbf{x}_k, \delta\mathbf{v}_k, \delta\mathbf{w}_{k+1}) \\ \tilde{\mathbf{y}}_{k+1}^{[2]} - \mathbf{y}_{k+1}^{[2]}(\delta\mathbf{x}_k, \delta\mathbf{v}_k, \delta\mathbf{w}_{k+1}) \\ \ldots \\ \tilde{\mathbf{y}}_{k+1}^{[\eta]} - \mathbf{y}_{k+1}^{[\eta]}(\delta\mathbf{x}_k, \delta\mathbf{v}_k, \delta\mathbf{w}_{k+1}) \end{bmatrix} \tag{40}$$

The updated distribution (polynomial) of the state is given by

$$\mathbf{x}_{k+1}^{+}(\delta\mathbf{x}_k, \delta\mathbf{v}_k, \delta\mathbf{w}_{k+1}) = \mathbf{x}_{k+1}^{-}(\delta\mathbf{x}_k, \delta\mathbf{v}_k) + \mathcal{K}d\tilde{\mathcal{Y}}(\delta\mathbf{x}_k, \delta\mathbf{v}_k, \delta\mathbf{w}_{k+1}) \tag{41}$$

and the posterior estimate is its mean

$$\hat{\mathbf{x}}_{k+1}^{+} = \mathbb{E}\left\{\mathbf{x}_{k+1}^{+}(\delta\mathbf{x}_k, \delta\mathbf{v}_k, \delta\mathbf{w}_{k+1})\right\} \tag{42}$$

evaluated, through Isserlis's moments, monomial by monomial using Table 1.

Equation (41) shows that the state polynomials are function of the three different deviations: the state deviation, the process noise, and the measurement noise. Furthermore, the new order of the polynomial as increased of a factor η, dictated by the order of the polynomial update. If the order of the polynomial approximation of the prior distribution ($\mathbf{x}_{k+1}^{-}(\delta\mathbf{x}_k, \delta\mathbf{v}_k)$) is c, then, the order of the posterior polynomial ($\mathbf{x}_{k+1}^{+}(\delta\mathbf{x}_k, \delta\mathbf{v}_k, \delta\mathbf{w}_{k+1})$) is ηc. The higher the polynomial order, the higher the number of moments to be calculated by Table 1, which leads to a higher computational burden for the machine.

The Covariance Polynomial Update

The third, and last, part of SACE-c-η-μ is the covariance polynomial update. After having estimated the state of the system, SACE-c-η-μ applies a second polynomial estimator to identify the value of the state covariance conditioned to the measurements. Therefore, the user defines one last integer parameter, μ, that specifies the order of the covariance polynomial update. Unlike previous tuning parameters, μ cannot be freely chosen but it has to respect the inequality $\mu < \eta$. The covariance cannot have an higher update order w.r.t. the state in order to achieve a correct matching.

The covariance matrix is obtained as

$$\mathbf{P}_{\mathbf{xx},k+1} = \mathbb{E}\left\{(\mathbf{x}_{k+1}^+ - \hat{\mathbf{x}}_{k+1}^+)(\mathbf{x}_{k+1}^+ - \hat{\mathbf{x}}_{k+1}^+)^{\mathrm{T}}\right\} \tag{43}$$

This value shows the average spread of the posterior distribution among all different possible realizations, $\tilde{\mathbf{y}}$, of the random variable $\mathbf{y}$. Equation (43) is the equivalent of the classical covariance update formulation, Equation (2), that is used in the most common filters such as EKF, UKF, QKF, CBF, Central Difference Filter, GSOF, etc. Therefore, even if correct, the error covariance mean only extracts the least possible information from the measurement outcome. Having in mind the polynomial estimator presented in Equation (15), Equation (43) can be intended as the first constant of the estimator, where all the correction terms from the Kalman gain are missing.

A new approach is therefore presented. Define a polynomial vector, $\boldsymbol{\rho}_{k+1}$, as the covariance polynomial

$$\boldsymbol{\rho}_{k+1}^-(\delta\mathbf{x}_k, \delta\mathbf{v}_k, \delta\mathbf{w}_{k+1}) = (\mathbf{x}_{k+1}^+ - \hat{\mathbf{x}}_{k+1}^+) \otimes (\mathbf{x}_{k+1}^+ - \hat{\mathbf{x}}_{k+1}^+) \tag{44}$$

where, in order to reduce the computational burden, the redundant terms have been eliminated. The covariance polynomial maximum order is $2\eta c$, being the square of the posterior distribution. The mean of $\boldsymbol{\rho}_{k+1}$ is exactly the vectorized version of the covariance matrix expressed in Equation (43)

$$\hat{\boldsymbol{\rho}}_{k+1}^- = \mathbb{E}\left\{\boldsymbol{\rho}_{k+1}^-(\delta\mathbf{x}_k, \delta\mathbf{v}_k, \delta\mathbf{w}_{k+1})\right\} = \text{stack}(\mathbf{P}_{\mathbf{xx},k+1}) \tag{45}$$

Thus, the covariance polynomial can be considered as the state of a new system where the prior is known. A second polynomial estimator, that uses the same measurement outcome $\tilde{\mathbf{y}}_{k+1}$, will provide a covariance value that better fits the state estimate of the system.

The augmented measurement covariance matrix $\mathbf{P}_{\mathcal{YY},[\mu]}$ has already being computed. Having the constrain $\mu < \eta$ makes $\mathbf{P}_{\mathcal{YY},[\mu]}$ just a subpart of $\mathbf{P}_{\mathcal{YY},[\eta]}$. On the other side, the covariance-measurement cross covariance matrix $\mathbf{P}_{\boldsymbol{\rho}\mathcal{Y},[\mu]}$ is evaluated blockwise according to Equation (16), where each block is obtained as

$$\mathbf{P}_{\boldsymbol{\rho}\mathbf{y}^{[i]}} = \mathbb{E}\left\{(\boldsymbol{\rho}_{k+1}^- - \hat{\boldsymbol{\rho}}_{k+1}^-)(\mathbf{y}_{k+1}^{[i]} - \hat{\mathbf{y}}_{k+1}^{[i]})^{\mathrm{T}}\right\} \tag{46}$$

with $i = 1, \ldots, \mu$. The Kalman gain for the covariance correction is calculated as

$$\mathcal{G} = \mathbf{P}_{\boldsymbol{\rho}\mathcal{Y},[\mu]}\mathbf{P}_{\mathcal{YY},[\mu]}^{-1} \tag{47}$$

The covariance is updated to its posterior estimate

$$\hat{\boldsymbol{\rho}}_{k+1}^+ = \hat{\boldsymbol{\rho}}_{k+1}^- + \mathcal{G}\begin{bmatrix} \tilde{\mathbf{y}}_{k+1} - \hat{\mathbf{y}}_{k+1} \\ \tilde{\mathbf{y}}_{k+1}^{[2]} - \hat{\mathbf{y}}_{k+1}^{[2]} \\ \ldots \\ \tilde{\mathbf{y}}_{k+1}^{[\mu]} - \hat{\mathbf{y}}_{k+1}^{[\mu]} \end{bmatrix} \tag{48}$$

where the influence of the measurement is weighted by the augmented Kalman gain. Before starting the next iteration, vector $\hat{\boldsymbol{\rho}}_{k+1}^+$ is brough back to its matrix formulation

$$\hat{\mathbf{P}}_{\mathbf{xx},k+1} = \text{matrix}(\hat{\boldsymbol{\rho}}_{k+1}^+) \tag{49}$$

9

The updated posterior distribution has been approximated as Gaussian with mean $\hat{\mathbf{x}}_{k+1}^{+}$ and covariance matrix $\hat{\mathbf{P}}_{\mathbf{xx},k+1}$. The filter is able to start the next iteration step from Equation (24), where the DA variables related to the noises are discarded and a new state deviation vector is initialized.

SACE-c-η-μ relies on its three tuning parameters and it enhances the performances of classic estimators. In fact, SACE-1-1-0 reduces to the extended Kalman filter and SACE-2-1-0 is the Gaussian Second Order Filter. The polynomial estimator better weights the information from the measurement by computing high order central moments. The increase in accuracy is paid by the increse of computational effort, which limits the filter's order selection. In fact, the highest polynomial order the filter has to compute (in the evaluation of $\mathbf{P}_{\boldsymbol{\rho}\mathbf{y}^{[\mu]}}$) is $(2\eta + \mu)c$.

NUMERICAL EXAMPLES

The proposed filtering techniques have been applied to two different scenarios. At first, a scalar application gives a visual representation on how the new update algorithm work and which are the innovative features when compared to other estimators. The second problem consists in a tracking application where the system undergoes the highly nonlinear dynamics of a Lorenz96 system.

Scalar Problem

A simple scalar problem is here presented to highlight the improvements of the new filtering technique by estimating the conditional covariance. It has already been proven that high order polynomial estimators are a better approximation of the true MMSE [27]. However, the presented example underlines the matching between state and covariance for each different realization of the measurement.

Define a normal prior state distribution $x \sim \mathcal{N}(1, 0.02)$ and a measurement

$$y = 1/x + \nu \tag{50}$$

where $\nu \sim \mathcal{N}(0, 0.003)$ is measurement noise. Figure 1 shows the true joint distribution of x and y represented using 10^5 points (gray dots in the figure). The figure compares the most common estimators, such as the EKF, the UKF, the GSF, and the high-order extended Kalman filter (DAHO-k) developed in [14] with SACE-c-η-μ. The first row of graphs (EKF, UKF, DAHO-3) is linear estimators, therefore their representation on the (x, y) plane is a straight line, in red. The slope of the red line is the Kalman gain, which optimal value is $\mathbf{P}_{\mathbf{xy}}\mathbf{P}_{\mathbf{yy}}^{-1}$. Therefore, the different slope is related to the different approximation used by the filters to evaluate moments. The EKF applies basic linearization (Jacobians), the UKF uses the unscented transformation, and DAHO-3 uses Taylor polynomials up to the third order. The purple lines express the estimated uncertainties as a $\pm 3\sigma$ boundary. The different evaluation of the moments leads to a different value on the covariance estimation, since it follows Equation (2). The purple lines share the same slopes of the red line: they are just translated left (and right) of 3σ. These linear filters estimate the same uncertainty level regardless the measurement outcome and the predicted covariance value is the mean among all the possible different realizations. The second row of graphs in Figure 1 shows nonlinear estimators. The GSF has been implemented with 3 models, which allows the estimator function, red line, to follow the curved shape of the posterior distribution. However, when the likelihood of one model becomes predominant with respect to the others, the GSF behaves similarly to the EKF: this aspect is mostly evident near the tails of the distribution. The estimated covariance of the GSF is a function of the measurement because it is evaluated as a weighted mean among all the models, which importance weight is based on their likelihood. However, the $\pm 3\sigma$ purple lines show the same problems connected with linear estimators: the lines are able to change slope when the models have approximately the same weight, otherwise they are straight. Furthermore, since the GSF can be intended as multiple EKFs with reduced subdomains, the filter shows the same behavior of the linear estimator at the edges of the posterior PFD. Lastly SACE-3-5-2 is reported. The 5th order polynomial estimator is able to follow the curved shape of the joint distribution and it better approximates the true MMSE. The optimal MMSE is the conditional mean, which visually is the line that divides in half the distribution of y, as horizontal spread of points, for each value of x. Therefore, while EKF, UKF and DAHO-3 can be
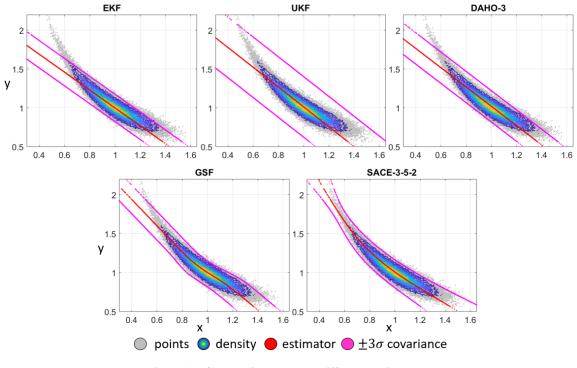
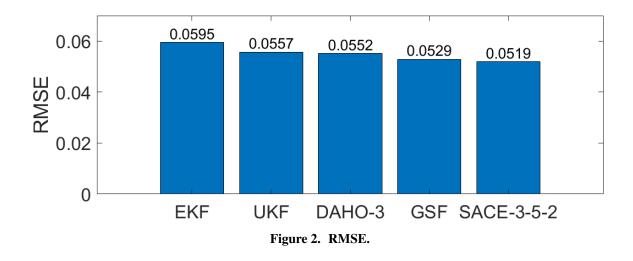Figure 1. Comparison among different estimators.

interpreted as different linear approximation of the true MMSE, SACE-3-5-2 represents a 5th order approximation, which shows a more accurate result. By increasing the estimator order η to infinity, SACE-c-η-μ would asymptotically reach the true MMSE. The purple lines related to SACE-3-5-2 show how the uncertainty level has become a (nonlinear) function of the measurement. The $\pm 3\sigma$ boundary increases and tightens depending on the horizontal spread of samples around the estimator function. For example, when the current measurement is $y = 1$, SACE-3-5-2 gives its estimate with a level of uncertainty that matches the spread of the gray points on the line $y = 1$. On the contrary, when the sensor gives $y = 2$, SACE-3-5-2 outputs a level of confidence in its estimate higher than in the previous case, since the spread of the gray samples around its estimate curve at $y = 2$ is tighter. Therefore, the estimated covariance of the filter is a function of the measurement and the performances improve drastically because the uncertainties level always matches the estimate, providing a more reliable outcome.

The accuracy level reached by each filter is compared in Figure 2, where the results of a RMSE analysis is reported.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N_{samples}}(x_i - \hat{x}_i^+)^2}{N_{samples}}} \tag{51}$$

The RMSE of each estimator is evaluated using the entire set of 10^5 points. The bars show that SACE-3-5-2 is the most accurate filter while the linear estimators are the least. However, a more precise approximation of the measurement equation leads to a smaller RMSE and to a more precise estimate, as proven by DAHO-3 (3rd order Taylor polynomial) being the most accurate among the other linear estimators.

The proposed problem underlines a couple of points. Unlike the linear and Gaussian case, the conditional covariance and the estimation error covariance are different. For the common filter, the estimation error variance expresses the average spread of the estimation error over all possible measurement realizations. This is a good metric for sure, but once a measurement is actually available to process, the conditional variance is a much more informative quantity, because it effectively give the spread of the estimation error for this particular value of y. In fact, the conditional covariance is a (nonlinear) function of the measurement which

11

Figure 2. RMSE.

evaluation is usually non feasible. However, SACE-c-η-μ uses a polynomial estimator to approximate the function, achieving better results with respect to filters that do not.

Lorenz96 System

The performance of the proposed filter is tested on a Lorenz96 example [21] where the state dynamics are

$$\frac{dx_i(t)}{dt} = x_{i-1}(t)(x_{i+1}(t) - x_{i-2}(t)) - x_i(t) + F + \nu_i(t) \tag{52}$$

with $i = 1, \ldots, 4$, since $\mathbf{x}(t)$ is selected to be four-dimensional. The following conventions are used: $x_{-1}(t) = x_{n-1}(t)$, $x_0(t) = x_n(t)$, and $x_1(t) = x_{n+1}(t)$. The term F is a constant external force which value is chosen equal to eight, since it introduces a chaotic behavior in the system. The initial condition is assumed to be Gaussian, with mean $\hat{\mathbf{x}} = \begin{bmatrix} F & F & F + 0.01 & F \end{bmatrix}^T$ and diagonal covariance matrix, with the same standard deviation for each component of the state: $\sigma_{\mathbf{x}} = 10^{-3}$. The process noise is assumed to be Gaussian and uncorrelated among states, with known standard deviation $\sigma_{\boldsymbol{\nu}} = 10^{-3}$. The dynamics are propagated at 2 Hz for a total of 20 seconds. The measurement are obtained each time step according to the following model

$$\mathbf{y}_k = \mathbf{H}_{i,j}\mathbf{x}(t_k) + \boldsymbol{\mu}_k, \quad \mathbf{H} = \begin{cases} 1 & j = 2i - 1 \\ 0 & \text{otherwise} \end{cases} \tag{53}$$

with $i = \{1, 2\}$ and $j = \{1, 2, 3, 4\}$. In other words, the sensors provide the components of the state with odd indeces. Measurement noises are assumed to be Gaussian and uncorrelated within each other and with the process noise. The standard deviation is selected as $\sigma_{\boldsymbol{\mu}} = 0.5$: this value is particularly high and filters based on linear estimators are not able to track the state of the system and achieve convergence [27].

Figure 3 shows the Monte Carlo analysis results performed with SACE-2-3-2 on the presented application. The figure shows, for each ith component of the state, the estimation error of each realization (gray lines), calculated as

$$\epsilon_{j,i} = x_{j,i} - \hat{x}_{j,i} \tag{54}$$

for each jth time step. A total of 100 realizations are reported. Figure 3 describes the error means, in black, and the error standard deviations, as a 3σ, in blue. The black lines show that SACE-c-η-μ is an unbiased filter, as expected from the theory of MMSE estimation. The predicted error standard deviation, continuous blue line, is evaluated directly from the updated covariance matrix, by taking the square root of the diagonal terms. The effective performance of the filter is assessed by the sample standard deviation of the Monte Carlo estimation errors, dashed blue lines. At each time step, the actual error covariance of the filter is evaluated
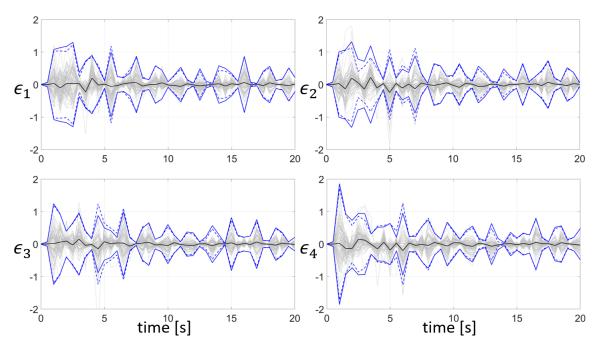
Figure 3. 100 runs Monte Carlo performance test for SACE-2-3-2.

by working directly on the samples. The consistency of SACE-2-3-2 is established by the overlapping of the dashed and continuous blue lines, which proves that the filter can correctly predict its own uncertainty levels.

The performance comparison among different filters is shown in Figure 4 through another Monte Carlo analysis conducted with 100 runs. The figure shows, for each filter, the comparison between the effective and predicted error covariance. The continuous lines represent the filter own estimate of the error standard deviation, calculated directly form the updated covariance matrix as the square root of its trace:

$$\bar{\sigma} = \sqrt{\text{tr}(\hat{\mathbf{P}}_{\mathbf{xx}})} \tag{55}$$

The dashed lines represent the effective error standard deviation derived from the Monte Carlo analysis. A consistent filter has the matching between its dashed and continuous line, meaning that the estimated uncertainty level reflects the actual error standard deviation. The top graph in Figure 4 shows how linear estimators, the EKF, UKF, and DAHO-2, diverge (and break down) while trying to track the state of the system. The measurement noise level is excessively large and a linear dependence on the measurement outcome is not sufficient to achieve a correct estimate. The UKF and DAHO-2 uses, respectively, the unscented transformation and second order Taylor polynomial to improve the prediction step of the filter and have a more accurate propagated state prior distribution. However, the update step is still linear and highly influenced by the noise standard deviation which prevents the evaluation of a reliable Kalman gain. The polynomial estimator better weights the information from the measurements using high order moments and it achieves convergence with consistency. Therefore, SACE-2-3-0, in blue, and SACE-2-3-2, in red, correctly estimate the state of the system along the whole simulation. The bottom graph in Figure 4 zooms in for the performance of SACE-c-η-μ for the two different sets of parameters. SACE-2-3-0 shows a filter which estimate is a polynomial function of the measurement and its estimated covariance is evaluated as a mean among all possible resolutions; it is not influenced by the measurement outcome. On the other side, SACE-2-3-2 improves accuracy by estimating the covariance giving it the same importance reserved to the state. Thus, the red lines settle below the blue ones for the whole simulation, since the predicted error standard deviation better matches the conditional mean.
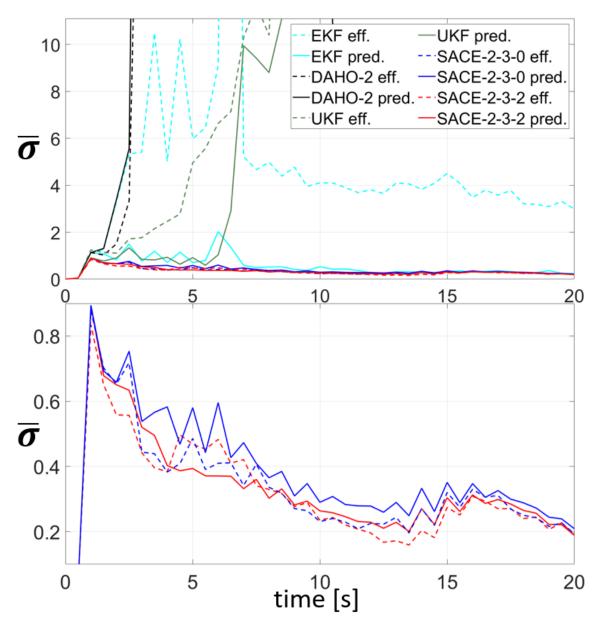
Figure 4. Covariance comparison among different filters.

CONCLUSIONS

A filter based on a double estimator has been presented. The new technique estimates the conditional mean and the conditional covariance of the posterior distribution by applying, sequentially, two polynomial estimators, using the same measurement outcome. The new approach better matches the estimated state with its error standard deviation, which is now a polynomial function of the measurement. Therefore, the newly proposed filter is able to reduce the error uncertainty when the posterior distribution gets narrower around a low probability realization of the measurement.

Two numerical examples have been reported. The scalar application gives a visual representation of the benefits of the polynomial approximation of the true MMSE and its covariance. Thus, the higher the order of the updates, the more precise the relative state estimate and its covariance. The vectorial application under-

lines the benefits of predicting the covariance by considering its estimation as working with an augmented state. The new state estimate improves in accuracy and a smaller error standard deviation is obtained.

ACKNOWLEDGMENT

REFERENCES

[1] R. E. Kalman, A new approach to linear filtering and prediction problems, Journal of Basic Engineering 82 (Series D) (1960) 35–45, doi:10.1115/1.3662552.

[2] R. E. Kalman, R. S. Bucy, New results in linear filtering and prediction, Journal of Basic Engineering 83 (Series D) (1961) 95–108, doi:10.1115/1.3658902.

[3] B. D. Tapley, B. E. Schutz, G. H. Born, Statistical Orbit Determination, Elsevier Academic Press, 2004, ISBN: 9780080541730.

[4] A. Gelb (Ed.), Applied Optimal Estimation, The MIT press, Cambridge, MA, 1974, iSBN:9780262200271.

[5] J. Junkins, P. Singla, How nonlinear is it? a tutorial on nonlinearity of orbit and attitude dynamics, Journal of the Astronautical Sciences 52 (1-2) (2004) 7–60.

[6] S. J. Julier, J. K. Uhlmann, Unscented filtering and nonlinear estimation, Proceedings of the IEEE 92 (3) (2004) 401–422, doi: 10.1109/JPROC.2003.823141.

[7] S. J. Julier, J. K. Uhlmann, H. F. Durrant-Whyte, A new method for the nonlinear transformation of means and covariances in filters and estimators, IEEE Transactions on Automatic Control 45 (3) (2000) 477–482, doi: 10.1109/9.847726.

[8] F. Cavenago, P. Di Lizia, M. Massari, S. Servadio, A. Wittig, Da-based nonlinear filters for spacecraft relative state estimation, in: 2018 Space Flight Mechanics Meeting, 2018, p. 1964, doi: 10.2514/6.2018-1964.

[9] A. H. Jazwinski, Stochastic Processes and Filtering Theory, Vol. 64 of Mathematics in Sciences and Engineering, Academic Press, New York, New York 10003, 1970, iSBN: 9780486318196.

[10] S. Servadio, High order filters for relative pose estimation of an uncooperative target, Master's thesis, Politecnico di Milano, Milan,Italy,20156 (2017).

[11] R. S. Park, D. J. Scheeres, Nonlinear mapping of gaussian statistics: theory and applications to spacecraft trajectory design, Journal of Guidance, Control, and Dynamics 29 (6) (2006) 1367–1375, doi: 10.2514/1.20177.

[12] R. S. Park, D. J. Scheeres, Nonlinear semi analytic methods for trajectory estimation, Journal of Guidance, Control, and Dynamics 30 (6) (2007) 1668–1676, doi: 10.2514/1.29106.

[13] M. Majji, J. Junkins, J. Turner, A high order method for estimation of dynamic systems, The Journal of the Astronautical Sciences 56 (3) (2008) 401–440, doi: 10.1007/BF03256560.

[14] M. Valli, R. Armellin, P. Di Lizia, M. Lavagna, Nonlinear mapping of uncertainties in celestial mechanics, Journal of Guidance, Control, and Dynamics 36 (1) (2012) 48–63, doi: 10.2514/1.58068.

[15] H. W. Sorenson, D. L. Alspach, Recursive Bayesian Estimation Using Gaussian Sums, Automatica 7 (4) (1971) 465–479, doi:10.1016/0005-1098(71)90097-5.

[16] D. Alspach, H. Sorenson, Nonlinear bayesian estimation using gaussian sum approximations, IEEE Transactions on Automatic Control 17 (4) (1972) 439–448, doi: 10.1109/SAP.1970.270017.

[17] N. J. Gordon, D. J. Salmond, A. F. Smith, Novel approach to nonlinear/non-gaussian bayesian state estimation, in: IEE proceedings F (radar and signal processing), Vol. 140, IET, 1993, pp. 107–113.

[18] J. H. Kotecha, P. M. Djuric, Gaussian sum particle filtering, IEEE Transactions on signal processing 51 (10) (2003) 2602–2612.

[19] G. Terejanu, P. Singla, T. Singh, P. D. Scott, Adaptive Gaussian Sum Filter for Nonlinear Bayesian Estimation, IEEE Transactions on Automatic Control 56 (9) (2011) 2151–2156.

[20] K. J. DeMars, R. H. Bishop, M. K. Jah, An entropy-based approach for uncertainty propagation of nonlinear dynamical systems, Journal of Guidance, Control, and Dynamics 36 (4) (2013) 1047–1057.

[21] D. Raihan, S. Chakravorty, Particle gaussian mixture filters-i, Automatica 98 (2018) 331–340.

[22] M. Valli, R. Armellin, P. Di Lizia, M. Lavagna, A gaussian particle filter based on differential algebra for spacecraft navigation.

[23] S. Servadio, R. Zanetti., Da-based multiple gaussians particle filter for orbit determination, in: AAS/AIAA - Astrodynamics Specialist Conference, 2020.

[24] A. De Santis, A. Germani, M. Raimondi, Optimal quadratic filtering of linear discrete-time non-gaussian systems, IEEE Transactions on Automatic Control 40 (7) (1995) 1274–1278. doi:doi:10.1109/9.400478.

[25] F. Carravetta, A. Germani, N. Raimondi, Polynomial filtering of discrete-time stochastic linear systems with multiplicative state noise, IEEE Transactions on Automatic Control 42 (8) (1997) 1106–1126, doi:10.1109/9.618240.

[26] S. Servadio, R. Zanetti, Recursive polynomial minimum mean-square error estimation with applications to orbit determination, Journal of Guidance, Control, and Dynamics 43 (5) (2020) 939–954.

[27] S. Servadio, R. Zanetti, B. A. Jones, Nonlinear filtering with polynomial series of gaussian random variables, in: Proceedings of the 23rd International Conference on Information Fusion, 2020.

[28] T. Lefebvre, H. Bruyninckx, J. D. Schutter, Comment on "A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators", IEEE Transactions on Automatic Control 47 (8) (2002) 1406–1408, doi: 10.1109/TAC.2002.800742.

[29] I. Arasaratnam, S. Haykin, R. J. Elliot, Discrete-time nonlinear filtering algorithms using gauss-hermite quadrature, Proceedings of the IEEE 95 (5) (2007) 953–977, doi: 10.1019/JPROC.2007.894705.

[30] I. Arasaratnam, S. Haykin, Cubature kalman filters, IEEE Transactions on automatic control 54 (6) (2009) 1254–1269.

[31] T. S. Schei, A finite-difference method for linearization in nonlinear estimation algorithms, Automatica 33 (1997) 2053–2058, doi: 10.1016/s005-1098(97)00127-1.

[32] O. Ravn, M. Norgaard, N. K. Poulsen, New developments in state estimations for nonlinear systems, Automatica 36 (11) (2000) 1627–1638, doi: 10.1016/s005-1098(00)000089-3.

[33] T. Schon, F. Gustafsson, P.-J. Nordlund, Marginalized particle filters for mixed linear/nonlinear state-space models, IEEE Transactions on signal processing 53 (7) (2005) 2279–2289.

[34] R. Van Der Merwe, A. Doucet, N. De Freitas, E. A. Wan, The unscented particle filter, in: Advances in neural information processing systems, 2001, pp. 584–590.

[35] M. Murata, H. Nagano, K. Kashino, Monte carlo filter particle filter, in: 2015 European Control Conference (ECC), IEEE, 2015, pp. 2836–2841.

[36] B. M., Modern Map Methods in Particle Beam Physics, Academic Press, 1999.

[37] P. Di Lizia, M. Massari, F. Cavenago, Assessment of onboard da state estimation for spacecraft relative navigationFinal report, ESA .

[38] M. Rasotto, A. Morselli, A. Wittig, M. Massari, P. Di Lizia, R. Armellin, C. Valles, G. Ortega, Differential algebra space toolbox for nonlinear uncertainty propagation in space dynamics.

[39] K. Makino, M. Berz, Cosy infinity version 9, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 558 (1) (2006) 346–350, doi: 10.1016/j.nima.2005.11.109.

[40] M. Massari, P. Di Lizia, F. Cavenago, A. Wittig, Differential algebra software library with automatic code generation for space embedded applications, in: 2018 AIAA Information Systems-AIAA Infotech@ Aerospace, 2018, p. 0398, doi: 10.2514/6.2018-0398.

[41] R. Armellin, P. Di Lizia, F. Bernelli-Zazzera, M. Berz, Asteroid close encounters characterization using differential algebra: the case of apophis, Celestial Mechanics and Dynamical Astronomy 107 (4) (2010) 451–470, doi: 10.1007/s10569-010-9283-5.

[42] R. Armellin, P. Di Lizia, Probabilistic optical and radar initial orbit determination, Journal of Guidance, Control, and Dynamics 41 (1) (2018) 101–118.

[43] S. Servadio, R. Zanetti, Maximum a posteriori estimation of hamiltonian systems with high order series expansions.

[44] A. Wittig, P. Di Lizia, R. Armellin, K. Makino, F. Bernelli-Zazzera, M. Berz, Propagation of large uncertainty sets in orbital dynamics by automatic domain splitting, Celestial Mechanics and Dynamical Astronomy 122 (3) (2015) 239–261.

[45] L. Isserlis, On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables, Biometrika 12 (1/2) (1918) 134–139.